

SEQUENCE LISTING

<110> Schimmel, Paul
Wakasugi, Keisuke

<120> Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
The Regulation of Angiogenesis

<130> 00-221

<140>

<141>

<160> 58

<170> PatentIn Ver. 2.0

<210> 1

<211> 5174

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

<222> (3428)..(5035)

<220>

<223> Description of Artificial Sequence: human
full-length TyrRS in pET20B

<400> 1

tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60
cagcgtgacc gctacacttg ccagcgccct agcgcccgt cctttcgtt tcttcccttc 120
ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc tcccttagg 180
gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc 240
acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 300
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggcttatct 360
ttttgattta taagggaatt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 420
acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480
tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540
tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat 600
gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttcctgt 660
ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg 720
agtgggttac atcgaactgg atctcaacag cggttaagatc cttgagagtt ttgcggccga 780
agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg 840
tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggt 900
tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg 960

cagtgtctgcc ataaccatga gtgataaac tgcggccaac ttactttctga caacgatcgg 1020
 aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga 1080
 tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc 1140
 tgcagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc 1200
 ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc 1260
 ggcccttcg gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg 1320
 cggatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac 1380
 gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga taggtgcctc 1440
 actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt 1500
 aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac 1560
 caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa 1620
 aggatcttct tgagatcctt ttttctgcg cgtaatctgc tgcttgcaa caaaaaacc 1680
 accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactctt ttccgaagg 1740
 aactggettc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg 1800
 ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa tcctgttacc 1860
 agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt 1920
 accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc ccagcttga 1980
 gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct 2040
 tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggctcgaa caggagagcg 2100
 cacgaggag cttccagggg gaaacgcctg gtatctttat agtctgtcg ggtttcgcca 2160
 cctctgactt gagcgtcgat ttttgtgatg ctgctcaggg gggcggagcc tatggaaaaa 2220
 cgccagcaac gcggcctttt tacggttcct ggcttttgc tggccttttg ctacatggt 2280
 ctttctgcg ttatcccctg attctgtgga taaccgtatt accgccttg agtgagctga 2340
 taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaga 2400
 gcgcctgatg cggtatcttc tccttacgca tctgtgcggt atttcacacc gcatatatgg 2460
 tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagtatac actccgctat 2520
 cgctacgtga ctgggtcatg gctgcgcccc gacaccgcc aacaccgct gacgcgcct 2580
 gacgggcttg tctgctccc gcatccgctt acagacaagc tgtgaccgtc tccgggagct 2640
 gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc gaggcagctg cggtaaagct 2700
 catcagcgtg gtcgtgaagc gattcacaga tgtctgcctg ttcacccg cccagctcgt 2760
 tgagtttctc cagaagcgtt aatgtctggc ttctgataaa gcgggccatg ttaagggcg 2820

ttttttcttg	tttggtcact	gatgcctccg	tgtaaggggg	atttctgttc	atggggggtaa	2880
tgataccgat	gaaacgagag	aggatgctca	cgatacgggt	tactgatgat	gaacatgccc	2940
ggttactgga	acgttgtgag	ggtaaacaac	tggcggtatg	gatgcggcgg	gaccagagaa	3000
aaatcactca	gggtcaatgc	cagcgcttcg	ttaatacaga	tgtaggtggt	ccacagggta	3060
gccagcagca	tcctgcgatg	cagatccgga	acataatggt	gcagggcgct	gacttccgcg	3120
tttccagact	ttacgaaaca	cggaaaccca	agaccattca	tgttgttgct	caggtcgcag	3180
acgttttgca	gcagcagtcg	cttcacgttc	gctcgcgtat	cggtgattca	ttctgctaac	3240
cagtaaggca	accccgccag	cctagccggg	tcctcaacga	caggagcacg	atcatgcgca	3300
cccgtggcca	ggacccaacg	ctgcccgaga	tctcgatccc	gcgaaattaa	tacgactcac	3360
tatagggaga	ccacaacggt	ttccctctag	aaataatttt	gtttaacttt	aagaaggaga	3420
tatacat	atg ggg gac gct ccc agc cct gaa gag aaa ctg cac ctt atc	3469				
	Met Gly Asp Ala Pro Ser Pro Glu Glu Lys Leu His Leu Ile					
	1 5 10					
acc cgg aac ctg cag gag gtt ctg ggg gaa gag aag ctg aag gag ata	3517					
Thr Arg Asn Leu Gln Val Leu Gly Glu Glu Lys Leu Lys Glu Ile						
15 20 25 30						
ctg aag gag cgg gaa ctt aaa att tac tgg gga acg gca acc acg ggc	3565					
Leu Lys Glu Arg Glu Leu Lys Ile Tyr Trp Gly Thr Ala Thr Thr Gly						
35 40 45						
aaa cca cat gtg gct tac ttt gtg ccc atg tca aag att gca gac ttc	3613					
Lys Pro His Val Ala Tyr Phe Val Pro Met Ser Lys Ile Ala Asp Phe						
50 55 60						
tta aag gca ggg tgt gag gta aca att ctg ttt gcg gac ctc cac gca	3661					
Leu Lys Ala Gly Cys Glu Val Thr Ile Leu Phe Ala Asp Leu His Ala						
65 70 75						
tac ctg gat aac atg aaa gcc cca tgg gaa ctt cta gaa ctc cga gtc	3709					
Tyr Leu Asp Asn Met Lys Ala Pro Trp Glu Leu Leu Glu Leu Arg Val						
80 85 90						
agt tac tat gag aat gtg atc aaa gca atg ctg gag agc att ggt gtg	3757					
Ser Tyr Tyr Glu Asn Val Ile Lys Ala Met Leu Glu Ser Ile Gly Val						
95 100 105 110						
ccc ttg gag aag ctc aag ttc atc aaa ggc act gat tac cag ctc agc	3805					
Pro Leu Glu Lys Leu Lys Phe Ile Lys Gly Thr Asp Tyr Gln Leu Ser						
115 120 125						
aaa gag tac aca cta gat gtg tac aga ctc tcc tcc gtg gtc aca cag	3853					
Lys Glu Tyr Thr Leu Asp Val Tyr Arg Leu Ser Ser Val Val Thr Gln						
130 135 140						
cac gat tcc aag aag gct gga gct gag gtg gta aag cag gtg gag cac	3901					
His Asp Ser Lys Lys Ala Gly Ala Glu Val Val Lys Gln Val Glu His						
145 150 155						
cct ttg ctg agt ggc ctc tta tac ccc gga ctg cag gct ttg gat gaa	3949					

Pro	Leu	Leu	Ser	Gly	Leu	Leu	Tyr	Pro	Gly	Leu	Gln	Ala	Leu	Asp	Glu		
160						165					170						
gag	tat	tta	aaa	gta	gat	gcc	caa	ttt	gga	ggc	att	gat	cag	aga	aag	3997	
Glu	Tyr	Leu	Lys	Val	Asp	Ala	Gln	Phe	Gly	Gly	Ile	Asp	Gln	Arg	Lys		
175					180					185					190		
att	ttc	acc	ttt	gca	gag	aag	tac	ctc	cct	gca	ctt	ggc	tat	tca	aaa	4045	
Ile	Phe	Thr	Phe	Ala	Glu	Lys	Tyr	Leu	Pro	Ala	Leu	Gly	Tyr	Ser	Lys		
				195					200					205			
cgg	gtc	cat	ctg	atg	aat	cct	atg	gtt	cca	gga	tta	aca	ggc	agc	aaa	4093	
Arg	Val	His	Leu	Met	Asn	Pro	Met	Val	Pro	Gly	Leu	Thr	Gly	Ser	Lys		
			210					215					220				
atg	agc	tct	tca	gaa	gag	gag	tcc	aag	att	gat	ctc	ctt	gat	cgg	aag	4141	
Met	Ser	Ser	Ser	Glu	Glu	Glu	Ser	Lys	Ile	Asp	Leu	Leu	Asp	Arg	Lys		
			225				230					235					
gag	gat	gtg	aag	aaa	aaa	ctg	aag	aag	gcc	ttc	tgt	gag	cca	gga	aat	4189	
Glu	Asp	Val	Lys	Lys	Lys	Leu	Lys	Lys	Ala	Phe	Cys	Glu	Pro	Gly	Asn		
	240					245					250						
gtg	gag	aac	aat	ggg	gtt	ctg	tcc	ttc	atc	aag	cat	gtc	ctt	ttt	ccc	4237	
Val	Glu	Asn	Asn	Gly	Val	Leu	Ser	Phe	Ile	Lys	His	Val	Leu	Phe	Pro		
255					260					265					270		
ctt	aag	tcc	gag	ttt	gtg	atc	cta	cga	gat	gag	aaa	tgg	ggt	gga	aac	4285	
Leu	Lys	Ser	Glu	Phe	Val	Ile	Leu	Arg	Asp	Glu	Lys	Trp	Gly	Gly	Asn		
				275					280					285			
aaa	acc	tac	aca	gct	tac	gtg	gac	ctg	gaa	aag	gac	ttt	gct	gct	gag	4333	
Lys	Thr	Tyr	Thr	Ala	Tyr	Val	Asp	Leu	Glu	Lys	Asp	Phe	Ala	Ala	Glu		
			290					295					300				
gtt	gta	cat	cct	gga	gac	ctg	aag	aat	tct	gtt	gaa	gtc	gca	ctg	aac	4381	
Val	Val	His	Pro	Gly	Asp	Leu	Lys	Asn	Ser	Val	Glu	Val	Ala	Leu	Asn		
		305				310						315					
aag	ttg	ctg	gat	cca	atc	cgg	gaa	aag	ttt	aat	acc	cct	gcc	ctg	aaa	4429	
Lys	Leu	Leu	Asp	Pro	Ile	Arg	Glu	Lys	Phe	Asn	Thr	Pro	Ala	Leu	Lys		
	320					325					330						
aaa	ctg	gcc	agc	gct	gcc	tac	cca	gat	ccc	tca	aag	cag	aag	cca	atg	4477	
Lys	Leu	Ala	Ser	Ala	Ala	Tyr	Pro	Asp	Pro	Ser	Lys	Gln	Lys	Pro	Met		
335					340					345					350		
gcc	aaa	ggc	cct	gcc	aag	aat	tca	gaa	cca	gag	gag	gtc	atc	cca	tcc	4525	
Ala	Lys	Gly	Pro	Ala	Lys	Asn	Ser	Glu	Pro	Glu	Glu	Val	Ile	Pro	Ser		
				355					360					365			
cgg	ctg	gat	atc	cgt	gtg	ggg	aaa	atc	atc	act	gtg	gag	aag	cac	cca	4573	
Arg	Leu	Asp	Ile	Arg	Val	Gly	Lys	Ile	Ile	Thr	Val	Glu	Lys	His	Pro		
			370				375						380				
gat	gca	gac	agc	ctg	tat	gta	gag	aag	att	gac	gtg	ggg	gaa	gct	gaa	4621	
Asp	Ala	Asp	Ser	Leu	Tyr	Val	Glu	Lys	Ile	Asp	Val	Gly	Glu	Ala	Glu		
			385			390						395					
cca	cgg	act	gtg	gtg	agc	ggc	ctg	gta	cag	ttc	gtg	ccc	aag	gag	gaa	4669	
Pro	Arg	Thr	Val	Val	Ser	Gly	Leu	Val	Gln	Phe	Val	Pro	Lys	Glu	Glu		
	400					405					410						

```

ctg cag gac agg ctg gta gtg gtg ctg tgc aac ctg aaa ccc cag aag 4717
Leu Gln Asp Arg Leu Val Val Val Leu Cys Asn Leu Lys Pro Gln Lys
415 420 425 430

atg aga gga gtc gag tcc caa ggc atg ctt ctg tgt gct tct ata gaa 4765
Met Arg Gly Val Glu Ser Gln Gly Met Leu Leu Cys Ala Ser Ile Glu
435 440 445

ggg ata aac cgc cag gtt gaa cct ctg gac cct ccg gca ggc tct gct 4813
Gly Ile Asn Arg Gln Val Glu Pro Leu Asp Pro Pro Ala Gly Ser Ala
450 455 460

cct ggt gag cac gtg ttt gtg aag ggc tat gaa aag ggc caa cca gat 4861
Pro Gly Glu His Val Phe Val Lys Gly Tyr Glu Lys Gly Gln Pro Asp
465 470 475

gag gag ctc aag ccc aag aag aaa gtc ttc gag aag ttg cag gct gac 4909
Glu Glu Leu Lys Pro Lys Lys Lys Val Phe Glu Lys Leu Gln Ala Asp
480 485 490

ttc aaa att tct gag gag tgc atc gca cag tgg aag caa acc aac ttc 4957
Phe Lys Ile Ser Glu Glu Cys Ile Ala Gln Trp Lys Gln Thr Asn Phe
495 500 505 510

atg acc aag ctg ggc tcc att tcc tgt aaa tcg ctg aaa ggg ggg aac 5005
Met Thr Lys Leu Gly Ser Ile Ser Cys Lys Ser Leu Lys Gly Gly Asn
515 520 525

att agc ctc gag cac cac cac cac cac cac tgagatccgg ctgctaacaa 5055
Ile Ser Leu Glu His His His His His His
530 535

agccccgaaag gaagctgagt tggctgctgc caccgctgag caataactag cataaccct 5115

tgggggcctct aaacgggtct tgagggggttt tttgctgaaa ggaggaacta tatccgat 5174

```

```

<210> 2
<211> 536
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: human
full-length TyrRS in pET20B

```

```

<400> 2
Met Gly Asp Ala Pro Ser Pro Glu Glu Lys Leu His Leu Ile Thr Arg
1 5 10 15

Asn Leu Gln Glu Val Leu Gly Glu Glu Lys Leu Lys Glu Ile Leu Lys
20 25 30

Glu Arg Glu Leu Lys Ile Tyr Trp Gly Thr Ala Thr Thr Gly Lys Pro
35 40 45

His Val Ala Tyr Phe Val Pro Met Ser Lys Ile Ala Asp Phe Leu Lys
50 55 60

```

Ala Gly Cys Glu Val Thr Ile Leu Phe Ala Asp Leu His Ala Tyr Leu
 65 70 75 80
 Asp Asn Met Lys Ala Pro Trp Glu Leu Leu Glu Leu Arg Val Ser Tyr
 85 90 95
 Tyr Glu Asn Val Ile Lys Ala Met Leu Glu Ser Ile Gly Val Pro Leu
 100 105 110
 Glu Lys Leu Lys Phe Ile Lys Gly Thr Asp Tyr Gln Leu Ser Lys Glu
 115 120 125
 Tyr Thr Leu Asp Val Tyr Arg Leu Ser Ser Val Val Thr Gln His Asp
 130 135 140
 Ser Lys Lys Ala Gly Ala Glu Val Val Lys Gln Val Glu His Pro Leu
 145 150 155 160
 Leu Ser Gly Leu Leu Tyr Pro Gly Leu Gln Ala Leu Asp Glu Glu Tyr
 165 170 175
 Leu Lys Val Asp Ala Gln Phe Gly Gly Ile Asp Gln Arg Lys Ile Phe
 180 185 190
 Thr Phe Ala Glu Lys Tyr Leu Pro Ala Leu Gly Tyr Ser Lys Arg Val
 195 200 205
 His Leu Met Asn Pro Met Val Pro Gly Leu Thr Gly Ser Lys Met Ser
 210 215 220
 Ser Ser Glu Glu Glu Ser Lys Ile Asp Leu Leu Asp Arg Lys Glu Asp
 225 230 235 240
 Val Lys Lys Lys Leu Lys Lys Ala Phe Cys Glu Pro Gly Asn Val Glu
 245 250 255
 Asn Asn Gly Val Leu Ser Phe Ile Lys His Val Leu Phe Pro Leu Lys
 260 265 270
 Ser Glu Phe Val Ile Leu Arg Asp Glu Lys Trp Gly Gly Asn Lys Thr
 275 280 285
 Tyr Thr Ala Tyr Val Asp Leu Glu Lys Asp Phe Ala Ala Glu Val Val
 290 295 300
 His Pro Gly Asp Leu Lys Asn Ser Val Glu Val Ala Leu Asn Lys Leu
 305 310 315 320
 Leu Asp Pro Ile Arg Glu Lys Phe Asn Thr Pro Ala Leu Lys Lys Leu
 325 330 335
 Ala Ser Ala Ala Tyr Pro Asp Pro Ser Lys Gln Lys Pro Met Ala Lys
 340 345 350
 Gly Pro Ala Lys Asn Ser Glu Pro Glu Glu Val Ile Pro Ser Arg Leu
 355 360 365
 Asp Ile Arg Val Gly Lys Ile Ile Thr Val Glu Lys His Pro Asp Ala
 370 375 380
 Asp Ser Leu Tyr Val Glu Lys Ile Asp Val Gly Glu Ala Glu Pro Arg
 385 390 395 400

Thr Val Val Ser Gly Leu Val Gln Phe Val Pro Lys Glu Glu Leu Gln
 405 410 415
 Asp Arg Leu Val Val Val Leu Cys Asn Leu Lys Pro Gln Lys Met Arg
 420 425 430
 Gly Val Glu Ser Gln Gly Met Leu Leu Cys Ala Ser Ile Glu Gly Ile
 435 440 445
 Asn Arg Gln Val Glu Pro Leu Asp Pro Pro Ala Gly Ser Ala Pro Gly
 450 455 460
 Glu His Val Phe Val Lys Gly Tyr Glu Lys Gly Gln Pro Asp Glu Glu
 465 470 475 480
 Leu Lys Pro Lys Lys Lys Val Phe Glu Lys Leu Gln Ala Asp Phe Lys
 485 490 495
 Ile Ser Glu Glu Cys Ile Ala Gln Trp Lys Gln Thr Asn Phe Met Thr
 500 505 510
 Lys Leu Gly Ser Ile Ser Cys Lys Ser Leu Lys Gly Gly Asn Ile Ser
 515 520 525
 Leu Glu His His His His His His
 530 535

<210> 3
 <211> 4682
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> CDS
 <222> (3428)..(4543)

<220>
 <223> Description of Artificial Sequence: human mini
 TyrRS in pET20B

<400> 3
 tggcgaatgg gacgcgcgcoct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60
 cagcgtgacc gctacacttg ccagcgcgcoct agcgcgcgct cctttcgctt tcttcccttc 120
 ctttctcgcc acgttcgcgc gctttccccg tcaagctcta aatcgggggc tccctttagg 180
 gttccgattt agtgctttac ggcaacctga ccccaaaaaa cttgattagg gtgatggttc 240
 acgtagtggg ccatcgccct gatagacggt ttttcgcgcoct ttgacgttgg agtccacggt 300
 ctttaatagt ggactcttgt tccaaactgg aacaacactc aacctatct cggctctattc 360
 ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 420
 acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480
 tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540
 tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat 600

gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttcctgt 660
ttttgctcac ccagaaacgc tggatgaaagt aaaagatgct gaagatcagt tgggtgcacg 720
agtgggttac atcgaactgg atctcaacag cggtgaagatc cttgagagtt ttcgccccga 780
agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcggt tattatcccc 840
tattgacgcc gggcaagagc aactcggctg ccgcatacac tattctcaga atgacttggt 900
tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg 960
cagtgtctgc ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg 1020
aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga 1080
tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc 1140
tgagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc 1200
ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgtc 1260
ggcccttcg gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg 1320
cggatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac 1380
gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga taggtgcctc 1440
actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt 1500
aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac 1560
caaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa 1620
aggatcttct tgagatcctt ttttctgctg cgtaatctgc tgcttgcaaa caaaaaaacc 1680
accgctacca gcggtggtt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt 1740
aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg 1800
ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa tctgttacc 1860
agtggctgct gccagtggtg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt 1920
accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc ccagcttgga 1980
gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct 2040
tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg 2100
cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca 2160
cctctgactt gagcgtcgat ttttgtgatg ctgcgcaggg gggcgagacc tatggaaaaa 2220
cgccagcaac gcggcctttt tacggttctt ggcccttttg tggccttttg ctacatggt 2280
ctttcctgct ttatccccctg attctgtgga taaccgtatt accgcctttg agtgagctga 2340
taccgctcgc cgcagccgaa cgaccgagcg cagcagatca gtgagcgagg aagcggaaga 2400
gcgcctgatg cggtatcttc tcttacgca tctgtgcggt atttcacacc gcatatatgg 2460

tgcaactctca	gtacaatctg	ctctgatgcc	gcatagttaa	gccaggtatac	actccgctat	2520										
cgctacgtga	ctgggtcatg	gctgcgcccc	gacacccgcc	aacacccgct	gacgcgccct	2580										
gacgggcttg	tctgctccc	gcattccgctt	acagacaagc	tgtgaccgtc	tccgggagct	2640										
gcatgtgtca	gagggttttca	ccgtcatcac	cgaaacgcgc	gaggcagctg	cggtaaaagct	2700										
catcagcgtg	gtcgtgaagc	gattcacaga	tgtctgcctg	ttcatccgcg	tccagctcgt	2760										
tgagtttctc	cagaagcggt	aatgtctggc	ttctgataaa	gcggggccatg	ttaagggcg	2820										
ttttttcctg	tttggtcact	gatgcctccg	tgtaaggggg	atttctgttc	atgggggtaa	2880										
tgataccgat	gaaacgagag	aggatgctca	cgatacgggt	tactgatgat	gaacatgcc	2940										
ggttactgga	acgttgtgag	ggtaaacaac	tggcgggatg	gatgcggcgg	gaccagagaa	3000										
aaatcactca	gggtcaatgc	cagcgcttcg	ttaatacaga	tgtaggtgtt	ccacagggta	3060										
gccagcagca	tcctgcgatg	cagatccgga	acataatggt	gcagggcgct	gacttccgcg	3120										
tttccagact	ttacgaaaca	cggaaaccga	agaccattca	tgttgttgct	caggtcgcag	3180										
acgttttgca	gcagcagtcg	cttcacgttc	gctcgcgtat	cggtgattca	ttctgctaac	3240										
cagtaaggca	accccgccag	cctagccggg	tcctcaacga	caggagcacg	atcatgcgca	3300										
cccgtggcca	ggacccaacg	ctgcccgaga	tctcgatccc	gcgaaattaa	tacgactcac	3360										
tatagggaga	ccacaacggt	ttccctctag	aaataatttt	gtttaacttt	aagaaggaga	3420										
tatacat	atg	ggg	gac	gct	ccc	agc	cct	gaa	gag	aaa	ctg	cac	ctt	atc	3469	
	Met	Gly	Asp	Ala	Pro	Ser	Pro	Glu	Glu	Lys	Leu	His	Leu	Ile		
	1				5					10						
acc	cgg	aac	ctg	cag	gag	gtt	ctg	ggg	gaa	gag	aag	ctg	aag	gag	ata	3517
Thr	Arg	Asn	Leu	Gln	Glu	Val	Leu	Gly	Glu	Glu	Lys	Leu	Lys	Glu	Ile	
	15				20					25					30	
ctg	aag	gag	cgg	gaa	ctt	aaa	att	tac	tgg	gga	acg	gca	acc	acg	ggc	3565
Leu	Lys	Glu	Arg	Glu	Leu	Lys	Ile	Tyr	Trp	Gly	Thr	Ala	Thr	Thr	Gly	
				35					40					45		
aaa	cca	cat	gtg	gct	tac	ttt	gtg	ccc	atg	tca	aag	att	gca	gac	ttc	3613
Lys	Pro	His	Val	Ala	Tyr	Phe	Val	Pro	Met	Ser	Lys	Ile	Ala	Asp	Phe	
				50				55					60			
tta	aag	gca	ggg	tgt	gag	gta	aca	att	ctg	ttt	gcg	gac	ctc	cac	gca	3661
Leu	Lys	Ala	Gly	Cys	Glu	Val	Thr	Ile	Leu	Phe	Ala	Asp	Leu	His	Ala	
		65					70					75				
tac	ctg	gat	aac	atg	aaa	gcc	cca	tgg	gaa	ctt	cta	gaa	ctc	cga	gtc	3709
Tyr	Leu	Asp	Asn	Met	Lys	Ala	Pro	Trp	Glu	Leu	Leu	Glu	Leu	Arg	Val	
		80				85					90					
agt	tac	tat	gag	aat	gtg	atc	aaa	gca	atg	ctg	gag	agc	att	ggg	gtg	3757
Ser	Tyr	Tyr	Glu	Asn	Val	Ile	Lys	Ala	Met	Leu	Glu	Ser	Ile	Gly	Val	
		95			100					105					110	

ccc ttg gag aag ctc aag ttc atc aaa ggc act gat tac cag ctc agc	3805
Pro Leu Glu Lys Leu Lys Phe Ile Lys Gly Thr Asp Tyr Gln Leu Ser	
115 120 125	
aaa gag tac aca cta gat gtg tac aga ctc tcc tcc gtg gtc aca cag	3853
Lys Glu Tyr Thr Leu Asp Val Tyr Arg Leu Ser Ser Val Val Thr Gln	
130 135 140	
cac gat tcc aag aag gct gga gct gag gtg gta aag cag gtg gag cac	3901
His Asp Ser Lys Lys Ala Gly Ala Glu Val Val Lys Gln Val Glu His	
145 150 155	
cct ttg ctg agt ggc ctc tta tac ccc gga ctg cag gct ttg gat gaa	3949
Pro Leu Leu Ser Gly Leu Leu Tyr Pro Gly Leu Gln Ala Leu Asp Glu	
160 165 170	
gag tat tta aaa gta gat gcc caa ttt gga ggc att gat cag aga aag	3997
Glu Tyr Leu Lys Val Asp Ala Gln Phe Gly Gly Ile Asp Gln Arg Lys	
175 180 185 190	
att ttc acc ttt gca gag aag tac ctc cct gca ctt ggc tat tca aaa	4045
Ile Phe Thr Phe Ala Glu Lys Tyr Leu Pro Ala Leu Gly Tyr Ser Lys	
195 200 205	
cgg gtc cat ctg atg aat cct atg gtt cca gga tta aca ggc agc aaa	4093
Arg Val His Leu Met Asn Pro Met Val Pro Gly Leu Thr Gly Ser Lys	
210 215 220	
atg agc tct tca gaa gag gag tcc aag att gat ctc ctt gat cgg aag	4141
Met Ser Ser Ser Glu Glu Glu Ser Lys Ile Asp Leu Leu Asp Arg Lys	
225 230 235	
gag gat gtg aag aaa aaa ctg aag aag gcc ttc tgt gag cca gga aat	4189
Glu Asp Val Lys Lys Lys Leu Lys Lys Ala Phe Cys Glu Pro Gly Asn	
240 245 250	
gtg gag aac aat ggg gtt ctg tcc ttc atc aag cat gtc ctt ttt ccc	4237
Val Glu Asn Asn Gly Val Leu Ser Phe Ile Lys His Val Leu Phe Pro	
255 260 265 270	
ctt aag tcc gag ttt gtg atc cta cga gat gag aaa tgg ggt gga aac	4285
Leu Lys Ser Glu Phe Val Ile Leu Arg Asp Glu Lys Trp Gly Gly Asn	
275 280 285	
aaa acc tac aca gct tac gtg gac ctg gaa aag gac ttt gct gct gag	4333
Lys Thr Tyr Thr Ala Tyr Val Asp Leu Glu Lys Asp Phe Ala Ala Glu	
290 295 300	
gtt gta cat cct gga gac ctg aag aat tct gtt gaa gtc gca ctg aac	4381
Val Val His Pro Gly Asp Leu Lys Asn Ser Val Glu Val Ala Leu Asn	
305 310 315	
aag ttg ctg gat cca atc cgg gaa aag ttt aat acc cct gcc ctg aaa	4429
Lys Leu Leu Asp Pro Ile Arg Glu Lys Phe Asn Thr Pro Ala Leu Lys	
320 325 330	
aaa ctg gcc agc gct gcc tac cca gat ccc tca aag cag aag cca atg	4477
Lys Leu Ala Ser Ala Ala Tyr Pro Asp Pro Ser Lys Gln Lys Pro Met	
335 340 345 350	

gcc aaa ggc cct gcc aag aat tca gaa cca gag gag gtc atc ctc gag 4525
 Ala Lys Gly Pro Ala Lys Asn Ser Glu Pro Glu Glu Val Ile Leu Glu
 355 360 365

cac cac cac cac cac cac tgagatccgg ctgctaacaa agcccgaag 4573
 His His His His His His
 370

gaagctgagt tggctgctgc caccgctgag caataactag cataaccctt tggggcctct 4633

aaacgggtct tgagggggttt tttgctgaaa ggaggaacta tatccggat 4682

<210> 4

<211> 372

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: human mini
 TyrRS in pET20B

<400> 4

Met Gly Asp Ala Pro Ser Pro Glu Glu Lys Leu His Leu Ile Thr Arg
 1 5 10 15

Asn Leu Gln Glu Val Leu Gly Glu Glu Lys Leu Lys Glu Ile Leu Lys
 20 25 30

Glu Arg Glu Leu Lys Ile Tyr Trp Gly Thr Ala Thr Thr Gly Lys Pro
 35 40 45

His Val Ala Tyr Phe Val Pro Met Ser Lys Ile Ala Asp Phe Leu Lys
 50 55 60

Ala Gly Cys Glu Val Thr Ile Leu Phe Ala Asp Leu His Ala Tyr Leu
 65 70 75 80

Asp Asn Met Lys Ala Pro Trp Glu Leu Leu Glu Leu Arg Val Ser Tyr
 85 90 95

Tyr Glu Asn Val Ile Lys Ala Met Leu Glu Ser Ile Gly Val Pro Leu
 100 105 110

Glu Lys Leu Lys Phe Ile Lys Gly Thr Asp Tyr Gln Leu Ser Lys Glu
 115 120 125

Tyr Thr Leu Asp Val Tyr Arg Leu Ser Ser Val Val Thr Gln His Asp
 130 135 140

Ser Lys Lys Ala Gly Ala Glu Val Val Lys Gln Val Glu His Pro Leu
 145 150 155 160

Leu Ser Gly Leu Leu Tyr Pro Gly Leu Gln Ala Leu Asp Glu Glu Tyr
 165 170 175

Leu Lys Val Asp Ala Gln Phe Gly Gly Ile Asp Gln Arg Lys Ile Phe
 180 185 190

Thr Phe Ala Glu Lys Tyr Leu Pro Ala Leu Gly Tyr Ser Lys Arg Val
 195 200 205

12/52

His Leu Met Asn Pro Met Val Pro Gly Leu Thr Gly Ser Lys Met Ser
210 215 220

Ser Ser Glu Glu Glu Ser Lys Ile Asp Leu Leu Asp Arg Lys Glu Asp
225 230 235 240

Val Lys Lys Lys Leu Lys Lys Ala Phe Cys Glu Pro Gly Asn Val Glu
245 250 255

Asn Asn Gly Val Leu Ser Phe Ile Lys His Val Leu Phe Pro Leu Lys
260 265 270

Ser Glu Phe Val Ile Leu Arg Asp Glu Lys Trp Gly Gly Asn Lys Thr
275 280 285

Tyr Thr Ala Tyr Val Asp Leu Glu Lys Asp Phe Ala Ala Glu Val Val
290 295 300

His Pro Gly Asp Leu Lys Asn Ser Val Glu Val Ala Leu Asn Lys Leu
305 310 315 320

Leu Asp Pro Ile Arg Glu Lys Phe Asn Thr Pro Ala Leu Lys Lys Leu
325 330 335

Ala Ser Ala Ala Tyr Pro Asp Pro Ser Lys Gln Lys Pro Met Ala Lys
340 345 350

Gly Pro Ala Lys Asn Ser Glu Pro Glu Glu Val Ile Leu Glu His His
355 360 365

His His His His
370

<210> 5
<211> 4100
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (3428)..(3961)

<220>
<223> Description of Artificial Sequence: human TyrRS
carboxyl-terminal domain in pET20B

<400> 5
tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60
cagcgtgacc gctacacttg ccagcgccct agcgcccgtc cctttcgctt tcttccttc 120
ctttctcgcc acgttcgccg gctttcccg tcaagctcta aatcgggggc tcccttagg 180
gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc 240
acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 300
ctttaatagt ggactcttgt tccaaactgg aacaacactc aacctatct cggctctattc 360
ttttgattta taagggattt tgccgatttc ggccatttgg ttaaaaaatg agctgattta 420

acaaaaat t aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480
tcgggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540
tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat 600
gagtattcaa catttcctg tgcaccttat tccctttttt gggcatttt gccttcctgt 660
ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg 720
agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt ttgcacctga 780
agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg 840
tattgacgcc gggcaagagc aactcggctg ccgcatacac tattctcaga atgacttggt 900
tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg 960
cagtgtctgcc ataaccatga gtgataaacac tgcggccaac ttacttctga caacgatcgg 1020
aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga 1080
tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc 1140
tgcagcaatg gcaacaacgt tgcgcaaact attaacctggc gaactactta ctctagcttc 1200
ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcctc 1260
ggcccttcog gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg 1320
cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac 1380
gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga taggtgcctc 1440
actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt 1500
aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac 1560
caaatccct taacgtgagt ttctgttcca ctgagcgtca gaccccgtag aaaagatcaa 1620
aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc 1680
accgctacca gcggtggtt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt 1740
aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg 1800
ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc 1860
agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt 1920
accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc ccagcttgga 1980
gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct 2040
tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg 2100
cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca 2160
cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa 2220
cgccagcaac ggggcctttt tacggttcct ggccctttgc tggccttttg ctacatggt 2280

ctttcctg	cg	ttatcc	ctg	attctgt	gga	taaccgt	tatt	accgcct	ttg	agt	gag	ctga	2340			
taccgct	cg	cgagcc	gaa	cgaccg	agc	gagc	gtca	gtgag	cgagg	aagc	gga	aaga	2400			
gcgcct	gat	cggtatt	tttc	tccttac	gca	tctgt	gcggt	atttcac	acc	gcata	tat	tgg	2460			
tgcact	tctca	gtacaat	ctg	ctctgat	gcc	gcata	gttaa	gccagt	tatac	actcc	gctat	2520				
cgctac	gtga	ctgggt	catg	gctgc	gcccc	gacaccc	gcc	aacaccc	gct	gacgc	gccct	2580				
gacggg	cttg	tctgct	cccc	gcatcc	gctt	acagaca	aagc	tgtgac	cgtc	tccgg	gagct	2640				
gcatgt	gtca	gaggttt	tca	ccgtcat	cac	cgaaac	gcgc	gaggc	agctg	cggtaa	aagct	2700				
catcag	cg	gtcgt	gaagc	gattcac	aga	tgtct	gcctg	ttcat	ccgcg	tccag	ctcgt	2760				
tgagtt	tctc	cagaag	cg	tt	aatgt	ctggc	ttctg	ataaa	gcggg	ccatg	ttaagg	gcg	2820			
tttttt	cctg	tttggt	cact	gatgc	ctccg	tgtaag	gggg	atttct	gttc	atggg	gggtaa	2880				
tgatac	cgat	gaaac	gagag	aggat	gctca	cgatac	gggt	tactga	tgat	gaacat	gccc	2940				
ggttact	gga	acgtt	gtgag	ggtaa	acaac	tggcg	g	gtatg	gcggc	g	gaccag	agaa	3000			
aaatcact	tca	gggtca	aatgc	cagcg	cttcg	ttaata	caga	tgtag	gtgtt	ccacag	gggta	3060				
gccagc	agca	tcctgc	gatg	cagat	ccgga	acataat	ggt	gcagg	gcgct	gactt	ccgcg	3120				
tttccag	act	ttacgaa	aca	cggaa	accga	agaccat	tca	tgtt	gttgc	caggt	cg	3180				
acgtttt	gca	gcagc	agtcg	cttcac	gttc	gctcg	cg	tat	cggtg	attca	ttctg	ctaac	3240			
cagtaag	gca	accccg	ccag	cctag	cggg	tcctca	acga	caggag	cacg	atcat	gcgca	3300				
cccgtg	gcca	ggaccca	acg	ctgccc	gaga	tctcgat	ccc	gcgaa	attaa	tacgact	cac	3360				
tataggg	gaga	ccacaac	gg	ttccct	ctag	aaataat	ttt	gtttaac	ttt	aagaag	gaga	3420				
tatacat	atg	cca	gag	gag	gtc	atc	cca	tcc	cg	ctg	gat	atc	cg	gtg	3469	
	Met	Pro	Glu	Glu	Val	Ile	Pro	Ser	Arg	Leu	Asp	Ile	Arg	Val		
	1				5					10						
ggg	aaa	atc	atc	act	gtg	gag	aag	cac	cca	gat	gca	gac	agc	ctg	tat	3517
Gly	Lys	Ile	Ile	Thr	Val	Glu	Lys	His	Pro	Asp	Ala	Asp	Ser	Leu	Tyr	
15					20					25					30	
gta	gag	aag	att	gac	gtg	ggg	gaa	gct	gaa	cca	cg	act	gtg	gtg	agc	3565
Val	Glu	Lys	Ile	Asp	Val	Gly	Glu	Ala	Glu	Pro	Arg	Thr	Val	Val	Ser	
				35					40					45		
ggc	ctg	gta	cag	ttc	gtg	ccc	aag	gag	gaa	ctg	cag	gac	agg	ctg	gta	3613
Gly	Leu	Val	Gln	Phe	Val	Pro	Lys	Glu	Glu	Leu	Gln	Asp	Arg	Leu	Val	
			50					55					60			
gtg	gtg	ctg	tgc	aac	ctg	aaa	ccc	cag	aag	atg	aga	gga	gtc	gag	tcc	3661
Val	Val	Leu	Cys	Asn	Leu	Lys	Pro	Gln	Lys	Met	Arg	Gly	Val	Glu	Ser	
		65					70					75				
caa	ggc	atg	ctt	ctg	tgt	gct	tct	ata	gaa	ggg	ata	aac	cg	cag	gtt	3709
Gln	Gly	Met	Leu	Leu	Cys	Ala	Ser	Ile	Glu	Gly	Ile	Asn	Arg	Gln	Val	
	80					85					90					

```

gaa cct ctg gac cct ccg gca ggc tct gct cct ggt gag cac gtg ttt 3757
Glu Pro Leu Asp Pro Pro Ala Gly Ser Ala Pro Gly Glu His Val Phe
 95 100 105 110

gtg aag ggc tat gaa aag ggc caa cca gat gag gag ctc aag ccc aag 3805
Val Lys Gly Tyr Glu Lys Gly Gln Pro Asp Glu Glu Leu Lys Pro Lys
115 120 125

aag aaa gtc ttc gag aag ttg cag gct gac ttc aaa att tct gag gag 3853
Lys Lys Val Phe Glu Lys Leu Gln Ala Asp Phe Lys Ile Ser Glu Glu
130 135 140

tgc atc gca cag tgg aag caa acc aac ttc atg acc aag ctg ggc tcc 3901
Cys Ile Ala Gln Trp Lys Gln Thr Asn Phe Met Thr Lys Leu Gly Ser
145 150 155

att tcc tgt aaa tcg ctg aaa ggg ggg aac att agc ctc gag cac cac 3949
Ile Ser Cys Lys Ser Leu Lys Gly Gly Asn Ile Ser Leu Glu His His
160 165 170

cac cac cac cac tgagatccgg ctgctaacaa agcccgaag gaagctgagt 4001
His His His His
175

tggtgctgc caccgctgag caataactag cataaccct tggggcctct aaacgggtct 4061

tgagggggttt tttgctgaaa ggaggaacta tatccgat 4100

<210> 6
<211> 178
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: human TyrRS
carboxyl-terminal domain in pET20B

<400> 6
Met Pro Glu Glu Val Ile Pro Ser Arg Leu Asp Ile Arg Val Gly Lys
 1 5 10 15

Ile Ile Thr Val Glu Lys His Pro Asp Ala Asp Ser Leu Tyr Val Glu
20 25 30

Lys Ile Asp Val Gly Glu Ala Glu Pro Arg Thr Val Val Ser Gly Leu
35 40 45

Val Gln Phe Val Pro Lys Glu Glu Leu Gln Asp Arg Leu Val Val Val
50 55 60

Leu Cys Asn Leu Lys Pro Gln Lys Met Arg Gly Val Glu Ser Gln Gly
65 70 75 80

Met Leu Leu Cys Ala Ser Ile Glu Gly Ile Asn Arg Gln Val Glu Pro
85 90 95

Leu Asp Pro Pro Ala Gly Ser Ala Pro Gly Glu His Val Phe Val Lys
100 105 110

Gly Tyr Glu Lys Gly Gln Pro Asp Glu Glu Leu Lys Pro Lys Lys Lys
115 120 125

```

Val Phe Glu Lys Leu Gln Ala Asp Phe Lys Ile Ser Glu Glu Cys Ile
 130 135 140

Ala Gln Trp Lys Gln Thr Asn Phe Met Thr Lys Leu Gly Ser Ile Ser
 145 150 155 160

Cys Lys Ser Leu Lys Gly Gly Asn Ile Ser Leu Glu His His His His
 165 170 175

His His

<210> 7
 <211> 4682
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> CDS
 <222> (3428)..(4543)

<220>
 <223> Description of Artificial Sequence: human mini
 TyrRS mutant in pET20B

<400> 7
 tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60
 cagcgtgacc gctacacttg ccagcgcctt agcgcgcgct cctttcgctt tcttcccttc 120
 ctttctcgcc acgttcgccc gctttccccg tcaagctcta aatcgggggc tccctttagg 180
 gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc 240
 acgtagtggg ccacgcgcct gatagacggt ttttcgccct ttgacgttgg agtccacggt 300
 ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggctctattc 360
 ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 420
 acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480
 tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540
 tccgctcatg agacaataac cctgataaat gttcaataa tattgaaaaa ggaagagtat 600
 gagtattcaa catttccttg tcgcccttat tccctttttt gcggcatttt gccttctctg 660
 ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg 720
 agtgggttac atcgaactgg atctcaacag cggttaagatc cttgagagtt ttcgccccga 780
 agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg 840
 tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggg 900
 tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg 960
 cagtgtctgcc ataaccatga gtgataaacac tgcggccaac ttacttctga caacgatcgg 1020

aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga 1080
 tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc 1140
 tgcagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc 1200
 ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc 1260
 ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg 1320
 cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac 1380
 gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga taggtgcctc 1440
 actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt 1500
 aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac 1560
 caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa 1620
 aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaa caaaaaaacc 1680
 accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt 1740
 aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg 1800
 ccaccacttc aagaactctg tagcacgcc tacatactc gctctgctaa tcctgttacc 1860
 agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt 1920
 accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc ccagcttggg 1980
 gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct 2040
 tccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg 2100
 cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca 2160
 cctctgactt gagcgtcgat ttttgtgatg ctgcgcaggg gggcggagcc tatggaaaaa 2220
 cgccagcaac gcggcctttt tacggttcct ggccctttgc tggccttttg ctacatgtt 2280
 ctttctgcg ttatccctg attctgtgga taaccgtatt accgcctttg agtgagctga 2340
 taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga 2400
 gcgcctgatg cggtattttc tccttacgca tctgtgcggt atttcacacc gcatatatgg 2460
 tgactctca gtacaatctg ctctgatgcc gcatagttaa gccagtatac actccgctat 2520
 cgctacgtga ctgggtcatg gctgcgcccc gacaccgcc aacaccgct gacgcgcct 2580
 gacgggcttg tctgctccc gcatccgctt acagacaagc tgtgaccgtc tccgggagct 2640
 gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc gaggcagctg cggtaaagct 2700
 catcagcgtg gtcgtgaagc gattcacaga tgtctgcctg ttcacccgcg tccagctcgt 2760
 tgagtttctc cagaagcgtt aatgtctggc ttctgataaa gcgggccatg ttaagggcgg 2820
 tttttctg tttggtcact gatgcctccg tgtaagggg atttctgttc atgggggtaa 2880

tgataaccgat	gaaacgagag	aggatgctca	cgatacgggt	tactgatgat	gaacatgccc	2940										
ggttactgga	acgttgtag	ggtaaacaac	tggcggtatg	gatgcggcg	gaccagagaa	3000										
aaatcactca	gggtcaatgc	cagcgcttcg	ttaatacaga	tgtaggtggt	ccacagggta	3060										
gccagcagca	tcctgcgatg	cagatccgga	acataatggt	gcagggcgct	gacttccgcg	3120										
tttccagact	ttacgaaaca	cggaaaccga	agaccattca	tgttggtgct	caggtcgcag	3180										
acgttttgca	gcagcagtcg	cttcacgttc	gctcgcgtat	cggtgattca	ttctgctaac	3240										
cagtaaggca	accccgccag	cctagccggg	tcctcaacga	caggagcacg	atcatgcgca	3300										
cccgtggcca	ggaccaaacg	ctgcccgaga	tctcgatccc	gcgaaaattaa	tacgactcac	3360										
tatagggaga	ccacaacggt	ttccctctag	aaataatttt	gtttaacttt	aagaaggaga	3420										
tatacat	atg	ggg	gac	gct	ccc	agc	cct	gaa	gag	aaa	ctg	cac	ctt	atc	3469	
	Met	Gly	Asp	Ala	Pro	Ser	Pro	Glu	Glu	Lys	Leu	His	Leu	Ile		
	1				5					10						
acc	cgg	aac	ctg	cag	gag	gtt	ctg	ggg	gaa	gag	aag	ctg	aag	gag	ata	3517
Thr	Arg	Asn	Leu	Gln	Glu	Val	Leu	Gly	Glu	Glu	Lys	Leu	Lys	Glu	Ile	
	15				20					25					30	
ctg	aag	gag	cgg	gaa	ctt	aaa	att	tac	tgg	gga	acg	gca	acc	acg	ggc	3565
Leu	Lys	Glu	Arg	Glu	Leu	Lys	Ile	Tyr	Trp	Gly	Thr	Ala	Thr	Thr	Gly	
				35					40					45		
aaa	cca	cat	gtg	gct	tac	ttt	gtg	ccc	atg	tca	aag	att	gca	gac	ttc	3613
Lys	Pro	His	Val	Ala	Tyr	Phe	Val	Pro	Met	Ser	Lys	Ile	Ala	Asp	Phe	
			50					55					60			
tta	aag	gca	ggg	tgt	gag	gta	aca	att	ctg	ttt	gcg	gac	ctc	cac	gca	3661
Leu	Lys	Ala	Gly	Cys	Glu	Val	Thr	Ile	Leu	Phe	Ala	Asp	Leu	His	Ala	
		65					70					75				
tac	ctg	gat	aac	atg	aaa	gcc	cca	tgg	gaa	ctt	cta	gaa	ctg	cag	gtc	3709
Tyr	Leu	Asp	Asn	Met	Lys	Ala	Pro	Trp	Glu	Leu	Leu	Glu	Leu	Gln	Val	
	80					85					90					
agt	tac	tat	gag	aat	gtg	atc	aaa	gca	atg	ctg	gag	agc	att	ggg	gtg	3757
Ser	Tyr	Tyr	Glu	Asn	Val	Ile	Lys	Ala	Met	Leu	Glu	Ser	Ile	Gly	Val	
	95				100					105					110	
ccc	ttg	gag	aag	ctc	aag	ttc	atc	aaa	ggc	act	gat	tac	cag	ctc	agc	3805
Pro	Leu	Glu	Lys	Leu	Lys	Phe	Ile	Lys	Gly	Thr	Asp	Tyr	Gln	Leu	Ser	
				115					120					125		
aaa	gag	tac	aca	cta	gat	gtg	tac	aga	ctc	tcc	tcc	gtg	gtc	aca	cag	3853
Lys	Glu	Tyr	Thr	Leu	Asp	Val	Tyr	Arg	Leu	Ser	Ser	Val	Val	Thr	Gln	
			130					135					140			
cac	gat	tcc	aag	aag	gct	gga	gct	gag	gtg	gta	aag	cag	gtg	gag	cac	3901
His	Asp	Ser	Lys	Lys	Ala	Gly	Ala	Glu	Val	Val	Lys	Gln	Val	Glu	His	
		145				150						155				
cct	ttg	ctg	agt	ggc	ctc	tta	tac	ccc	gga	ctg	cag	gct	ttg	gat	gaa	3949
Pro	Leu	Leu	Ser	Gly	Leu	Leu	Tyr	Pro	Gly	Leu	Gln	Ala	Leu	Asp	Glu	
	160					165					170					

gag tat tta aaa gta gat gcc caa ttt gga ggc att gat cag aga aag Glu Tyr Leu Lys Val Asp Ala Gln Phe Gly Gly Ile Asp Gln Arg Lys 175 180 185 190	3997
att ttc acc ttt gca gag aag tac ctc cct gca ctt ggc tat tca aaa Ile Phe Thr Phe Ala Glu Lys Tyr Leu Pro Ala Leu Gly Tyr Ser Lys 195 200 205	4045
cgg gtc cat ctg atg aat cct atg gtt cca gga tta aca ggc agc aaa Arg Val His Leu Met Asn Pro Met Val Pro Gly Leu Thr Gly Ser Lys 210 215 220	4093
atg agc tct tca gaa gag gag tcc aag att gat ctc ctt gat cgg aag Met Ser Ser Ser Glu Glu Glu Ser Lys Ile Asp Leu Leu Asp Arg Lys 225 230 235	4141
gag gat gtg aag aaa aaa ctg aag aag gcc ttc tgt gag cca gga aat Glu Asp Val Lys Lys Lys Leu Lys Lys Ala Phe Cys Glu Pro Gly Asn 240 245 250	4189
gtg gag aac aat ggg gtt ctg tcc ttc atc aag cat gtc ctt ttt ccc Val Glu Asn Asn Gly Val Leu Ser Phe Ile Lys His Val Leu Phe Pro 255 260 265 270	4237
ctt aag tcc gag ttt gtg atc cta cga gat gag aaa tgg ggt gga aac Leu Lys Ser Glu Phe Val Ile Leu Arg Asp Glu Lys Trp Gly Gly Asn 275 280 285	4285
aaa acc tac aca gct tac gtg gac ctg gaa aag gac ttt gct gct gag Lys Thr Tyr Thr Ala Tyr Val Asp Leu Glu Lys Asp Phe Ala Ala Glu 290 295 300	4333
gtt gta cat cct gga gac ctg aag aat tct gtt gaa gtc gca ctg aac Val Val His Pro Gly Asp Leu Lys Asn Ser Val Glu Val Ala Leu Asn 305 310 315	4381
aag ttg ctg gat cca atc cgg gaa aag ttt aat acc cct gcc ctg aaa Lys Leu Leu Asp Pro Ile Arg Glu Lys Phe Asn Thr Pro Ala Leu Lys 320 325 330	4429
aaa ctg gcc agc gct gcc tac cca gat ccc tca aag cag aag cca atg Lys Leu Ala Ser Ala Ala Tyr Pro Asp Pro Ser Lys Gln Lys Pro Met 335 340 345 350	4477
gcc aaa ggc cct gcc aag aat tca gaa cca gag gag gtc atc ctc gag Ala Lys Gly Pro Ala Lys Asn Ser Glu Pro Glu Glu Val Ile Leu Glu 355 360 365	4525
cac cac cac cac cac cac tgagatccgg ctgctaacaa agccccgaaag His His His His His His 370	4573
gaagctgagt tggctgctgc caccgctgag caataactag cataaccct tggggcctct	4633
aaacgggtct tgaggggttt ttgctgaaa ggaggaacta tatccggat	4682

<210> 8
 <211> 372
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: human mini
 TyrRS mutant in pET20B

<400> 8

```

Met Gly Asp Ala Pro Ser Pro Glu Glu Lys Leu His Leu Ile Thr Arg
 1           5           10           15

Asn Leu Gln Glu Val Leu Gly Glu Glu Lys Leu Lys Glu Ile Leu Lys
      20           25           30

Glu Arg Glu Leu Lys Ile Tyr Trp Gly Thr Ala Thr Thr Gly Lys Pro
      35           40           45

His Val Ala Tyr Phe Val Pro Met Ser Lys Ile Ala Asp Phe Leu Lys
      50           55           60

Ala Gly Cys Glu Val Thr Ile Leu Phe Ala Asp Leu His Ala Tyr Leu
      65           70           75           80

Asp Asn Met Lys Ala Pro Trp Glu Leu Leu Glu Leu Gln Val Ser Tyr
      85           90           95

Tyr Glu Asn Val Ile Lys Ala Met Leu Glu Ser Ile Gly Val Pro Leu
      100          105          110

Glu Lys Leu Lys Phe Ile Lys Gly Thr Asp Tyr Gln Leu Ser Lys Glu
      115          120          125

Tyr Thr Leu Asp Val Tyr Arg Leu Ser Ser Val Val Thr Gln His Asp
      130          135          140

Ser Lys Lys Ala Gly Ala Glu Val Val Lys Gln Val Glu His Pro Leu
      145          150          155          160

Leu Ser Gly Leu Leu Tyr Pro Gly Leu Gln Ala Leu Asp Glu Glu Tyr
      165          170          175

Leu Lys Val Asp Ala Gln Phe Gly Gly Ile Asp Gln Arg Lys Ile Phe
      180          185          190

Thr Phe Ala Glu Lys Tyr Leu Pro Ala Leu Gly Tyr Ser Lys Arg Val
      195          200          205

His Leu Met Asn Pro Met Val Pro Gly Leu Thr Gly Ser Lys Met Ser
      210          215          220

Ser Ser Glu Glu Glu Ser Lys Ile Asp Leu Leu Asp Arg Lys Glu Asp
      225          230          235          240

Val Lys Lys Lys Leu Lys Lys Ala Phe Cys Glu Pro Gly Asn Val Glu
      245          250          255

Asn Asn Gly Val Leu Ser Phe Ile Lys His Val Leu Phe Pro Leu Lys
      260          265          270

```

Ser Glu Phe Val Ile Leu Arg Asp Glu Lys Trp Gly Gly Asn Lys Thr
 275 280 285

Tyr Thr Ala Tyr Val Asp Leu Glu Lys Asp Phe Ala Ala Glu Val Val
 290 295 300

His Pro Gly Asp Leu Lys Asn Ser Val Glu Val Ala Leu Asn Lys Leu
 305 310 315 320

Leu Asp Pro Ile Arg Glu Lys Phe Asn Thr Pro Ala Leu Lys Lys Leu
 325 330 335

Ala Ser Ala Ala Tyr Pro Asp Pro Ser Lys Gln Lys Pro Met Ala Lys
 340 345 350

Gly Pro Ala Lys Asn Ser Glu Pro Glu Glu Val Ile Leu Glu His His
 355 360 365

His His His His
 370

<210> 9
 <211> 5018
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> CDS
 <222> (3428)..(4879)

<220>
 <223> Description of Artificial Sequence: human
 full-length TrpRS in pET20B

<400> 9
 tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60
 cagcgtgacc gctacacttg ccagcgccct agcgcccgtt ctttcgctt tcttccttc 120
 ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc tccctttagg 180
 gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc 240
 acgtagtggg ccacgcgcct gatagacggt ttttcgcctt ttgacgttgg agtccacggt 300
 ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggctctattc 360
 ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 420
 acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480
 tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540
 tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat 600
 gagtattcaa catttccttg tcgcccttat tccctttttt gcggcatttt gccttcctgt 660
 ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg 720
 agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt ttcgccccga 780

agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcggt tattatcccg 840
 tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggg 900
 tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg 960
 cagtgtctgcc ataaccatga gtgataaacac tgcggccaac ttactttctga caacgatcgg 1020
 aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga 1080
 tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc 1140
 tgcagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc 1200
 ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc 1260
 ggcccttccg gctggctggg ttattgctga taaatctgga gccggtgagc gtgggtctcg 1320
 cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac 1380
 gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga taggtgcctc 1440
 actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt 1500
 aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac 1560
 caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa 1620
 aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc 1680
 accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt 1740
 aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg 1800
 ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc 1860
 agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt 1920
 accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttggg 1980
 gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct 2040
 tcccgaaggg agaaaggcgg acaggatatc ggtaagcggc agggtcggaa caggagagcg 2100
 cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca 2160
 cctctgactt gagcgtcgat ttttgtgatg ctctcaggg gggcggagcc tatggaaaaa 2220
 cgccagcaac gcggcctttt tacggttcct ggcccttttg tggccttttg ctacatggt 2280
 ctttctgctg ttatcccctg attctgtgga taaccgtatt accgccttg agtgagctga 2340
 taccgctcgc gcgagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga 2400
 gcgcctgatg cggatatttc tccttacgca tctgtgcggg atttcacacc gcatatatgg 2460
 tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagtatac actccgctat 2520
 cgctacgtga ctgggtcatg gctgcgcccc gacaccgcc aacaccgct gacgcgccct 2580
 gacgggcttg tctgctcccg gcatccgctt acagacaagc tgtgaccgtc tccgggagct 2640

gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc gaggcagctg cggtaaagct 2700
 catcagcgtg gtcgtgaagc gattcacaga tgtctgcctg ttcattccgcg tccagctcgt 2760
 tgagttttctc cagaagcgtt aatgtctggc ttctgataaa gcggggccatg ttaagggcgg 2820
 ttttttcctg tttggtcact gatgcctccg tgtaaggggg atttctgttc atgggggtaa 2880
 tgataccgat gaaacgagag aggatgctca cgatacgggt tactgatgat gaacatgccc 2940
 gggtactgga acgttgtgag ggtaaacaac tggcggtatg gatgcggcgg gaccagagaa 3000
 aaatcactca ggggtcaatgc cagcgccttcg ttaatacaga tgtaggtgtt ccacagggta 3060
 gccagcagca tctctgcgat cagatccgga acataatggt gcagggcgct gacttccgcg 3120
 tttccagact ttacgaaaca cggaaaccga agaccattca tgttggtgct caggtcgcag 3180
 acgttttgca gcagcagtcg cttcacgttc gctcgcgtat cggtgattca ttctgctaac 3240
 cagtaaggca accccgccag cctagccggg tctcaacga caggagcacg atcatgcgca 3300
 cccgtggcca ggacccaacg ctgcccgaga tctcgatccc gcgaaattaa tacgactcac 3360
 tatagggaga ccacaacggt ttccctctag aaataatttt gtttaacttt aagaaggaga 3420
 tatacat atg ccc aac agt gag ccc gca tct ctg ctg gag ctg ttc aac 3469
 Met Pro Asn Ser Glu Pro Ala Ser Leu Leu Glu Leu Phe Asn
 1 5 10
 agc atc gcc aca caa ggg gag ctc gta agg tcc ctc aaa gcg gga aat 3517
 Ser Ile Ala Thr Gln Gly Glu Leu Val Arg Ser Leu Lys Ala Gly Asn
 15 20 25 30
 gcg tca aag gat gaa att gat tct gca gta aag atg ttg gtg tca tta 3565
 Ala Ser Lys Asp Glu Ile Asp Ser Ala Val Lys Met Leu Val Ser Leu
 35 40 45
 aaa atg agc tac aaa gct gcc gcg ggg gag gat tac aag gct gac tgt 3613
 Lys Met Ser Tyr Lys Ala Ala Ala Gly Glu Asp Tyr Lys Ala Asp Cys
 50 55 60
 cct cca ggg aac cca gca cct acc agt aat cat ggc cca gat gcc aca 3661
 Pro Pro Gly Asn Pro Ala Pro Thr Ser Asn His Gly Pro Asp Ala Thr
 65 70 75
 gaa gct gaa gag gat ttt gtg gac cca tgg aca gta cag aca agc agt 3709
 Glu Ala Glu Glu Asp Phe Val Asp Pro Trp Thr Val Gln Thr Ser Ser
 80 85 90
 gca aaa ggc ata gac tac gat aag ctc att gtt cgg ttt gga agt agt 3757
 Ala Lys Gly Ile Asp Tyr Asp Lys Leu Ile Val Arg Phe Gly Ser Ser
 95 100 105 110
 aaa att gac aaa gag cta ata aac cga ata gag aga gcc acc ggc caa 3805
 Lys Ile Asp Lys Glu Leu Ile Asn Arg Ile Glu Arg Ala Thr Gly Gln
 115 120 125
 aga cca cac cac ttc ctg cgc aga ggc atc ttc ttc tca cac aga gat 3853
 Arg Pro His His Phe Leu Arg Arg Gly Ile Phe Phe Ser His Arg Asp
 130 135 140

atg aat cag gtt ctt gat gcc tat gaa aat aag aag cca ttt tat ctg	3901
Met Asn Gln Val Leu Asp Ala Tyr Glu Asn Lys Lys Pro Phe Tyr Leu	
145 150 155	
tac acg ggc cgg ggc ccc tct tct gaa gca atg cat gta ggt cac ctc	3949
Tyr Thr Gly Arg Gly Pro Ser Ser Glu Ala Met His Val Gly His Leu	
160 165 170	
att cca ttt att ttc aca aag tgg ctc cag gat gta ttt aac gtg ccc	3997
Ile Pro Phe Ile Phe Thr Lys Trp Leu Gln Asp Val Phe Asn Val Pro	
175 180 185 190	
ttg gtc atc cag atg acg gat gac gag aag tat ctg tgg aag gac ctg	4045
Leu Val Ile Gln Met Thr Asp Asp Glu Lys Tyr Leu Trp Lys Asp Leu	
195 200 205	
acc ctg gac cag gcc tat ggc gat gct gtt gag aat gcc aag gac atc	4093
Thr Leu Asp Gln Ala Tyr Gly Asp Ala Val Glu Asn Ala Lys Asp Ile	
210 215 220	
atc gcc tgt ggc ttt gac atc aac aag act ttc ata ttc tct gac ctg	4141
Ile Ala Cys Gly Phe Asp Ile Asn Lys Thr Phe Ile Phe Ser Asp Leu	
225 230 235	
gac tac atg ggg atg agc tca ggt ttc tac aaa aat gtg gtg aag att	4189
Asp Tyr Met Gly Met Ser Ser Gly Phe Tyr Lys Asn Val Val Lys Ile	
240 245 250	
caa aag cat gtt acc ttc aac caa gtg aaa ggc att ttc ggc ttc act	4237
Gln Lys His Val Thr Phe Asn Gln Val Lys Gly Ile Phe Gly Phe Thr	
255 260 265 270	
gac agc gac tgc att ggg aag atc agt ttt cct gcc atc cag gct gct	4285
Asp Ser Asp Cys Ile Gly Lys Ile Ser Phe Pro Ala Ile Gln Ala Ala	
275 280 285	
ccc tcc ttc agc aac tca ttc cca cag atc ttc cga gac agg acg gat	4333
Pro Ser Phe Ser Asn Ser Phe Pro Gln Ile Phe Arg Asp Arg Thr Asp	
290 295 300	
atc cag tgc ctt atc cca tgt gcc att gac cag gat cct tac ttt aga	4381
Ile Gln Cys Leu Ile Pro Cys Ala Ile Asp Gln Asp Pro Tyr Phe Arg	
305 310 315	
atg aca agg gac gtc gcc ccc agg atc ggc tat cct aaa cca gcc ctg	4429
Met Thr Arg Asp Val Ala Pro Arg Ile Gly Tyr Pro Lys Pro Ala Leu	
320 325 330	
ttg cac tcc acc ttc ttc cca gcc ctg cag ggc gcc cag acc aaa atg	4477
Leu His Ser Thr Phe Phe Pro Ala Leu Gln Gly Ala Gln Thr Lys Met	
335 340 345 350	
agt gcc agc gac cca aac tcc tcc atc ttc ctc acc gac acg gcc aag	4525
Ser Ala Ser Asp Pro Asn Ser Ser Ile Phe Leu Thr Asp Thr Ala Lys	
355 360 365	
cag atc aaa acc aag gtc aat aag cat gcg ttt tct gga ggg aga gac	4573
Gln Ile Lys Thr Lys Val Asn Lys His Ala Phe Ser Gly Gly Arg Asp	
370 375 380	


```

acc atc gag gag cac agg cag ttt ggg ggc aac tgt gat gtg gac gtg 4621
Thr Ile Glu Glu His Arg Gln Phe Gly Gly Asn Cys Asp Val Asp Val
      385              390              395

tct ttc atg tac ctg acc ttc ttc ctc gag gac gac gac aag ctc gag 4669
Ser Phe Met Tyr Leu Thr Phe Phe Leu Glu Asp Asp Asp Lys Leu Glu
      400              405              410

cag atc agg aag gat tac acc agc gga gcc atg ctc acc ggt gag ctc 4717
Gln Ile Arg Lys Asp Tyr Thr Ser Gly Ala Met Leu Thr Gly Glu Leu
      415              420              425              430

aag aag gca ctc ata gag gtt ctg cag ccc ttg atc gca gag cac cag 4765
Lys Lys Ala Leu Ile Glu Val Leu Gln Pro Leu Ile Ala Glu His Gln
              435              440              445

gcc cgg cgc aag gag gtc acg gat gag ata gtg aaa gag ttc atg act 4813
Ala Arg Arg Lys Glu Val Thr Asp Glu Ile Val Lys Glu Phe Met Thr
              450              455              460

ccc cgg aag ctg tcc ttc gac ttt cag aag ctt gcg gcc gca ctc gag 4861
Pro Arg Lys Leu Ser Phe Asp Phe Gln Lys Leu Ala Ala Ala Leu Glu
              465              470              475

cac cac cac cac cac cac tgagatccgg ctgctaacaa agcccgaaag 4909
His His His His His His
      480

gaagctgagt tggctgctgc caccgctgag caataactag cataaccocct tggggcctct 4969

aaacgggtct tgagggggttt tttgctgaaa ggaggaacta tatccggat 5018

<210> 10
<211> 484
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: human
      full-length TrpRS in pET20B

<400> 10
Met Pro Asn Ser Glu Pro Ala Ser Leu Leu Glu Leu Phe Asn Ser Ile
  1              5              10              15

Ala Thr Gln Gly Glu Leu Val Arg Ser Leu Lys Ala Gly Asn Ala Ser
      20              25              30

Lys Asp Glu Ile Asp Ser Ala Val Lys Met Leu Val Ser Leu Lys Met
      35              40              45

Ser Tyr Lys Ala Ala Ala Gly Glu Asp Tyr Lys Ala Asp Cys Pro Pro
      50              55              60

Gly Asn Pro Ala Pro Thr Ser Asn His Gly Pro Asp Ala Thr Glu Ala
      65              70              75              80

Glu Glu Asp Phe Val Asp Pro Trp Thr Val Gln Thr Ser Ser Ala Lys
      85              90              95

```

Gly Ile Asp Tyr Asp Lys Leu Ile Val Arg Phe Gly Ser Ser Lys Ile
 100 105 110
 Asp Lys Glu Leu Ile Asn Arg Ile Glu Arg Ala Thr Gly Gln Arg Pro
 115 120 125
 His His Phe Leu Arg Arg Gly Ile Phe Phe Ser His Arg Asp Met Asn
 130 135 140
 Gln Val Leu Asp Ala Tyr Glu Asn Lys Lys Pro Phe Tyr Leu Tyr Thr
 145 150 155 160
 Gly Arg Gly Pro Ser Ser Glu Ala Met His Val Gly His Leu Ile Pro
 165 170 175
 Phe Ile Phe Thr Lys Trp Leu Gln Asp Val Phe Asn Val Pro Leu Val
 180 185 190
 Ile Gln Met Thr Asp Asp Glu Lys Tyr Leu Trp Lys Asp Leu Thr Leu
 195 200 205
 Asp Gln Ala Tyr Gly Asp Ala Val Glu Asn Ala Lys Asp Ile Ile Ala
 210 215 220
 Cys Gly Phe Asp Ile Asn Lys Thr Phe Ile Phe Ser Asp Leu Asp Tyr
 225 230 235 240
 Met Gly Met Ser Ser Gly Phe Tyr Lys Asn Val Val Lys Ile Gln Lys
 245 250 255
 His Val Thr Phe Asn Gln Val Lys Gly Ile Phe Gly Phe Thr Asp Ser
 260 265 270
 Asp Cys Ile Gly Lys Ile Ser Phe Pro Ala Ile Gln Ala Ala Pro Ser
 275 280 285
 Phe Ser Asn Ser Phe Pro Gln Ile Phe Arg Asp Arg Thr Asp Ile Gln
 290 295 300
 Cys Leu Ile Pro Cys Ala Ile Asp Gln Asp Pro Tyr Phe Arg Met Thr
 305 310 315 320
 Arg Asp Val Ala Pro Arg Ile Gly Tyr Pro Lys Pro Ala Leu Leu His
 325 330 335
 Ser Thr Phe Phe Pro Ala Leu Gln Gly Ala Gln Thr Lys Met Ser Ala
 340 345 350
 Ser Asp Pro Asn Ser Ser Ile Phe Leu Thr Asp Thr Ala Lys Gln Ile
 355 360 365
 Lys Thr Lys Val Asn Lys His Ala Phe Ser Gly Gly Arg Asp Thr Ile
 370 375 380
 Glu Glu His Arg Gln Phe Gly Gly Asn Cys Asp Val Asp Val Ser Phe
 385 390 395 400
 Met Tyr Leu Thr Phe Phe Leu Glu Asp Asp Asp Lys Leu Glu Gln Ile
 405 410 415
 Arg Lys Asp Tyr Thr Ser Gly Ala Met Leu Thr Gly Glu Leu Lys Lys
 420 425 430

Ala Leu Ile Glu Val Leu Gln Pro Leu Ile Ala Glu His Gln Ala Arg
 435 440 445

Arg Lys Glu Val Thr Asp Glu Ile Val Lys Glu Phe Met Thr Pro Arg
 450 455 460

Lys Leu Ser Phe Asp Phe Gln Lys Leu Ala Ala Ala Leu Glu His His
 465 470 475 480

His His His His

<210> 11
 <211> 4877
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> CDS
 <222> (3428)..(4738)

<220>
 <223> Description of Artificial Sequence: human mini
 TrpRS in pET20B

<400> 11
 tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60
 cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc 120
 ctttctcgcc acgttcgccg gctttcccg tcaagctcta aatcgggggc tccctttagg 180
 gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatgggtc 240
 acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 300
 ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggctctattc 360
 ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 420
 acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480
 tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540
 tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat 600
 gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttcctgt 660
 ttttgctcac ccagaaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg 720
 agtgggttac atcgaactgg atctcaacag cggttaagatc cttgagagtt ttcgccccga 780
 agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg 840
 tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggt 900
 tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg 960
 cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga caacgatogg 1020

aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga 1080
 tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc 1140
 tgcagcaatg gcaacaacgt tgcgcaaact attaaactggc gaactactta ctctagcttc 1200
 ccggcaacaa ttaatagact ggatggaggg ggataaagtt gcaggaccac ttctgcgctc 1260
 ggcccttccg gctggctggg ttattgctga taaatctgga gccggtgagc gtgggtctcg 1320
 cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac 1380
 gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga taggtgcctc 1440
 actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt 1500
 aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac 1560
 caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa 1620
 aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc 1680
 accgctacca gcggtgggtt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt 1740
 aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg 1800
 ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa tcctgttacc 1860
 agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt 1920
 accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttggg 1980
 gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct 2040
 tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg 2100
 cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca 2160
 cctctgactt gagcgtcgat ttttgtgatg ctgcgcaggg gggcggagcc tatggaaaaa 2220
 cgccagcaac gcggcctttt tacggttctt ggcccttttg tggccttttg ctacatggtt 2280
 ctttctgctg ttatccctg attctgtgga taaccgtatt accgcctttg agtgagctga 2340
 taccgctcgc cgcagccgaa cgaccgagcg cagcagatca gtgagcgagg aagcggaaga 2400
 gcgcctgatg cggatatttc tccttacgca tctgtgcggg atttcacacc gcatatatgg 2460
 tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagtatac actccgctat 2520
 cgctacgtga ctgggtcatg gctgcgcccc gacaccgccc aacaccgct gagcgcctt 2580
 gacgggcttg tctgctcccg gcatccgctt acagacaagc tgtgaccgtc tccgggagct 2640
 gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc gaggcagctg cggtaaagct 2700
 catcagcgtg gtcgtgaagc gattcacaga tgtctgcctg ttcacccgctg tccagctcgt 2760
 tgagtttctc cagaagcgtt aatgtctggc ttctgataaa gcgggccatg ttaagggcgg 2820
 ttttttctg tttggtcact gatgcctccg tgtaaggggg atttctgttc atgggggtaa 2880

tatacat atg agc tac aaa gct gcc gcg ggg gag gat tac aag gct gac 3469
Met Ser Tyr Lys Ala Ala Ala Gly Glu Asp Tyr Lys Ala Asp
1 5 10

aca gaa gct gaa gag gat ttt gtg gac cca tgg aca gta cag aca agc 3565
Thr Glu Ala Glu Glu Asp Phe Val Asp Pro Trp Thr Val Gln Thr Ser
35 40 45

agt aaa att gac aaa gag cta ata aac cga ata gag aga gcc acc ggc 3661
Ser Lys Ile Asp Lys Glu Leu Ile Asn Arg Ile Glu Arg Ala Thr Gly
 65 70 75

gat	atg	aat	cag	gtt	ctt	gat	gcc	tat	gaa	aat	aag	aag	cca	ttt	tat	3757
Asp	Met	Asn	Gln	Val	Leu	Asp	Ala	Tyr	Glu	Asn	Lys	Lys	Pro	Phe	Tyr	
95					100					105					110	

ctc att cca ttt att ttc aca aag tgg ctc cag gat gta ttt aac gtg 3853
Leu Ile Pro Phe Ile Phe Thr Lys Trp Leu Gln Asp Val Phe Asn Val
130 135 140

ctg acc ctg gac cag gcc tat ggc gat gct gtt gag aat gcc aag gac 3949
Leu Thr Leu Asp Gln Ala Tyr Gly Asp Ala Val Glu Asn Ala Lys Asp
160 165 170

atc atc gcc tgt ggc ttt gac atc aac aag act ttc ata ttc tct gac Ile Ile Ala Cys Gly Phe Asp Ile Asn Lys Thr Phe Ile Phe Ser Asp 175 180 185 190	3997
ctg gac tac atg ggg atg agc tca ggt ttc tac aaa aat gtg gtg aag Leu Asp Tyr Met Gly Met Ser Ser Gly Phe Tyr Lys Asn Val Val Lys 195 200 205	4045
att caa aag cat gtt acc ttc aac caa gtg aaa ggc att ttc ggc ttc Ile Gln Lys His Val Thr Phe Asn Gln Val Lys Gly Ile Phe Gly Phe 210 215 220	4093
act gac agc gac tgc att ggg aag atc agt ttt cct gcc atc cag gct Thr Asp Ser Asp Cys Ile Gly Lys Ile Ser Phe Pro Ala Ile Gln Ala 225 230 235	4141
gct ccc tcc ttc agc aac tca ttc cca cag atc ttc cga gac agg acg Ala Pro Ser Phe Ser Asn Ser Phe Pro Gln Ile Phe Arg Asp Arg Thr 240 245 250	4189
gat atc cag tgc ctt atc cca tgt gcc att gac cag gat cct tac ttt Asp Ile Gln Cys Leu Ile Pro Cys Ala Ile Asp Gln Asp Pro Tyr Phe 255 260 265 270	4237
aga atg aca agg gac gtc gcc ccc agg atc ggc tat cct aaa cca gcc Arg Met Thr Arg Asp Val Ala Pro Arg Ile Gly Tyr Pro Lys Pro Ala 275 280 285	4285
ctg ttg cac tcc acc ttc ttc cca gcc ctg cag ggc gcc cag acc aaa Leu Leu His Ser Thr Phe Phe Pro Ala Leu Gln Gly Ala Gln Thr Lys 290 295 300	4333
atg agt gcc agc gac cca aac tcc tcc atc ttc ctc acc gac acg gcc Met Ser Ala Ser Asp Pro Asn Ser Ser Ile Phe Leu Thr Asp Thr Ala 305 310 315	4381
aag cag atc aaa acc aag gtc aat aag cat gcg ttt tct gga ggg aga Lys Gln Ile Lys Thr Lys Val Asn Lys His Ala Phe Ser Gly Gly Arg 320 325 330	4429
gac acc atc gag gag cac agg cag ttt ggg ggc aac tgt gat gtg gac Asp Thr Ile Glu Glu His Arg Gln Phe Gly Gly Asn Cys Asp Val Asp 335 340 345 350	4477
gtg tct ttc atg tac ctg acc ttc ttc ctc gag gac gac gac aag ctc Val Ser Phe Met Tyr Leu Thr Phe Phe Leu Glu Asp Asp Asp Lys Leu 355 360 365	4525
gag cag atc agg aag gat tac acc agc gga gcc atg ctc acc ggt gag Glu Gln Ile Arg Lys Asp Tyr Thr Ser Gly Ala Met Leu Thr Gly Glu 370 375 380	4573
ctc aag aag gca ctc ata gag gtt ctg cag ccc ttg atc gca gag cac Leu Lys Lys Ala Leu Ile Glu Val Leu Gln Pro Leu Ile Ala Glu His 385 390 395	4621
cag gcc cgg cgc aag gag gtc acg gat gag ata gtg aaa gag ttc atg Gln Ala Arg Arg Lys Glu Val Thr Asp Glu Ile Val Lys Glu Phe Met 400 405 410	4669

act ccc cgg aag ctg tcc ttc gac ttt cag aag ctt gcg gcc gca ctc 4717
 Thr Pro Arg Lys Leu Ser Phe Asp Phe Gln Lys Leu Ala Ala Ala Leu
 415 420 425 430

gag cac cac cac cac cac cac tgagatccgg ctgctaacaa agcccgaag 4768
 Glu His His His His His His
 435

gaagctgagt tggctgctgc caccgctgag caataactag cataaccct tggggcctct 4828

aaacgggtct tgaggggttt tttgctgaaa ggaggaacta tatccggat 4877

<210> 12

<211> 437

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: human mini
 TrpRS in pET20B

<400> 12

Met Ser Tyr Lys Ala Ala Ala Gly Glu Asp Tyr Lys Ala Asp Cys Pro
 1 5 10 15

Pro Gly Asn Pro Ala Pro Thr Ser Asn His Gly Pro Asp Ala Thr Glu
 20 25 30

Ala Glu Glu Asp Phe Val Asp Pro Trp Thr Val Gln Thr Ser Ser Ala
 35 40 45

Lys Gly Ile Asp Tyr Asp Lys Leu Ile Val Arg Phe Gly Ser Ser Lys
 50 55 60

Ile Asp Lys Glu Leu Ile Asn Arg Ile Glu Arg Ala Thr Gly Gln Arg
 65 70 75 80

Pro His His Phe Leu Arg Arg Gly Ile Phe Phe Ser His Arg Asp Met
 85 90 95

Asn Gln Val Leu Asp Ala Tyr Glu Asn Lys Lys Pro Phe Tyr Leu Tyr
 100 105 110

Thr Gly Arg Gly Pro Ser Ser Glu Ala Met His Val Gly His Leu Ile
 115 120 125

Pro Phe Ile Phe Thr Lys Trp Leu Gln Asp Val Phe Asn Val Pro Leu
 130 135 140

Val Ile Gln Met Thr Asp Asp Glu Lys Tyr Leu Trp Lys Asp Leu Thr
 145 150 155 160

Leu Asp Gln Ala Tyr Gly Asp Ala Val Glu Asn Ala Lys Asp Ile Ile
 165 170 175

Ala Cys Gly Phe Asp Ile Asn Lys Thr Phe Ile Phe Ser Asp Leu Asp
 180 185 190

Tyr Met Gly Met Ser Ser Gly Phe Tyr Lys Asn Val Val Lys Ile Gln
 195 200 205

Lys His Val Thr Phe Asn Gln Val Lys Gly Ile Phe Gly Phe Thr Asp
 210 215 220
 Ser Asp Cys Ile Gly Lys Ile Ser Phe Pro Ala Ile Gln Ala Ala Pro
 225 230 235 240
 Ser Phe Ser Asn Ser Phe Pro Gln Ile Phe Arg Asp Arg Thr Asp Ile
 245 250 255
 Gln Cys Leu Ile Pro Cys Ala Ile Asp Gln Asp Pro Tyr Phe Arg Met
 260 265 270
 Thr Arg Asp Val Ala Pro Arg Ile Gly Tyr Pro Lys Pro Ala Leu Leu
 275 280 285
 His Ser Thr Phe Phe Pro Ala Leu Gln Gly Ala Gln Thr Lys Met Ser
 290 295 300
 Ala Ser Asp Pro Asn Ser Ser Ile Phe Leu Thr Asp Thr Ala Lys Gln
 305 310 315 320
 Ile Lys Thr Lys Val Asn Lys His Ala Phe Ser Gly Gly Arg Asp Thr
 325 330 335
 Ile Glu Glu His Arg Gln Phe Gly Gly Asn Cys Asp Val Asp Val Ser
 340 345 350
 Phe Met Tyr Leu Thr Phe Phe Leu Glu Asp Asp Asp Lys Leu Glu Gln
 355 360 365
 Ile Arg Lys Asp Tyr Thr Ser Gly Ala Met Leu Thr Gly Glu Leu Lys
 370 375 380
 Lys Ala Leu Ile Glu Val Leu Gln Pro Leu Ile Ala Glu His Gln Ala
 385 390 395 400
 Arg Arg Lys Glu Val Thr Asp Glu Ile Val Lys Glu Phe Met Thr Pro
 405 410 415
 Arg Lys Leu Ser Phe Asp Phe Gln Lys Leu Ala Ala Ala Leu Glu His
 420 425 430
 His His His His His
 435

<210> 13
 <211> 4811
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> CDS
 <222> (3428)..(4672)

<220>
 <223> Description of Artificial Sequence: human
 supermini TrpRS in pET20B

<400> 13
 tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60

cagcgtgacc gctacacttg ccagcgcctt agcgcgcgtt cctttcgctt tcttcccttc 120
 ctttctogcc acgttcgccc gctttccccg tcaagctcta aatcgggggc tccctttagg 180
 gttccgattt agtgctttac ggcaacctga ccccaaaaaa cttgattagg gtgatggttc 240
 acgtagtggg ccatacgcct gatagacggt ttttcgcctt ttgacgttgg agtccacggt 300
 ctttaatagt ggactcttgt tccaaactgg aacaacactc aacctatctt cggctctattc 360
 ttttgattta taagggattt tgccgatttc ggccatttgg ttaaaaaatg agctgattta 420
 acaaaaaattt aacgcgaatt ttaacaaaaa attaacgttt acaatttcag gtggcacttt 480
 tcgggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540
 tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat 600
 gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttctgt 660
 ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg 720
 agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt ttcccccga 780
 agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg 840
 tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttgg 900
 tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg 960
 cagtgtgccc ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg 1020
 aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga 1080
 tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc 1140
 tgcagcaatg gcaacaacgt tgcgcaaact attaaactggc gaactactta ctctagcttc 1200
 ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgtc 1260
 ggcccttccg gctggctggg ttattgctga taaatctgga gccggtgagc gtgggtctcg 1320
 cggatcattt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac 1380
 gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga taggtgcctc 1440
 actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt 1500
 aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac 1560
 caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa 1620
 aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc 1680
 accgtacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt 1740
 aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg 1800
 ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa tcctgttacc 1860
 agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt 1920

[illegible]

cta ata aac cga ata gag aga gcc acc ggc caa aga cca cac cac ttc	3613
Leu Ile Asn Arg Ile Glu Arg Ala Thr Gly Gln Arg Pro His His Phe	
50 55 60	
ctg cgc aga ggc atc ttc ttc tca cac aga gat atg aat cag gtt ctt	3661
Leu Arg Arg Gly Ile Phe Phe Ser His Arg Asp Met Asn Gln Val Leu	
65 70 75	
gat gcc tat gaa aat aag aag cca ttt tat ctg tac acg ggc cgg ggc	3709
Asp Ala Tyr Glu Asn Lys Lys Pro Phe Tyr Leu Tyr Thr Gly Arg Gly	
80 85 90	
ccc tct tct gaa gca atg cat gta ggt cac ctc att cca ttt att ttc	3757
Pro Ser Ser Glu Ala Met His Val Gly His Leu Ile Pro Phe Ile Phe	
95 100 105 110	
aca aag tgg ctc cag gat gta ttt aac gtg ccc ttg gtc atc cag atg	3805
Thr Lys Trp Leu Gln Asp Val Phe Asn Val Pro Leu Val Ile Gln Met	
115 120 125	
acg gat gac gag aag tat ctg tgg aag gac ctg acc ctg gac cag gcc	3853
Thr Asp Asp Glu Lys Tyr Leu Trp Lys Asp Leu Thr Leu Asp Gln Ala	
130 135 140	
tat ggc gat gct gtt gag aat gcc aag gac atc atc gcc tgt ggc ttt	3901
Tyr Gly Asp Ala Val Glu Asn Ala Lys Asp Ile Ile Ala Cys Gly Phe	
145 150 155	
gac atc aac aag act ttc ata ttc tct gac ctg gac tac atg ggg atg	3949
Asp Ile Asn Lys Thr Phe Ile Phe Ser Asp Leu Asp Tyr Met Gly Met	
160 165 170	
agc tca ggt ttc tac aaa aat gtg gtg aag att caa aag cat gtt acc	3997
Ser Ser Gly Phe Tyr Lys Asn Val Val Lys Ile Gln Lys His Val Thr	
175 180 185 190	
ttc aac caa gtg aaa ggc att ttc ggc ttc act gac agc gac tgc att	4045
Phe Asn Gln Val Lys Gly Ile Phe Gly Phe Thr Asp Ser Asp Cys Ile	
195 200 205	
ggg aag atc agt ttt cct gcc atc cag gct gct ccc tcc ttc agc aac	4093
Gly Lys Ile Ser Phe Pro Ala Ile Gln Ala Ala Pro Ser Phe Ser Asn	
210 215 220	
tca ttc cca cag atc ttc cga gac agg acg gat atc cag tgc ctt atc	4141
Ser Phe Pro Gln Ile Phe Arg Asp Arg Thr Asp Ile Gln Cys Leu Ile	
225 230 235	
cca tgt gcc att gac cag gat cct tac ttt aga atg aca agg gac gtc	4189
Pro Cys Ala Ile Asp Gln Asp Pro Tyr Phe Arg Met Thr Arg Asp Val	
240 245 250	
gcc ccc agg atc ggc tat cct aaa cca gcc ctg ttg cac tcc acc ttc	4237
Ala Pro Arg Ile Gly Tyr Pro Lys Pro Ala Leu Leu His Ser Thr Phe	
255 260 265 270	
ttc cca gcc ctg cag ggc gcc cag acc aaa atg agt gcc agc gac cca	4285
Phe Pro Ala Leu Gln Gly Ala Gln Thr Lys Met Ser Ala Ser Asp Pro	
275 280 285	

aac tcc tcc atc ttc ctc acc gac acg gcc aag cag atc aaa acc aag 4333
 Asn Ser Ser Ile Phe Leu Thr Asp Thr Ala Lys Gln Ile Lys Thr Lys
 290 295 300

gtc aat aag cat gcg ttt tct gga ggg aga gac acc atc gag gag cac 4381
 Val Asn Lys His Ala Phe Ser Gly Gly Arg Asp Thr Ile Glu Glu His
 305 310 315

agg cag ttt ggg ggc aac tgt gat gtg gac gtg tct ttc atg tac ctg 4429
 Arg Gln Phe Gly Gly Asn Cys Asp Val Asp Val Ser Phe Met Tyr Leu
 320 325 330

acc ttc ttc ctc gag gac gac gac aag ctc gag cag atc agg aag gat 4477
 Thr Phe Phe Leu Glu Asp Asp Asp Lys Leu Glu Gln Ile Arg Lys Asp
 335 340 345 350

tac acc agc gga gcc atg ctc acc ggt gag ctc aag aag gca ctc ata 4525
 Tyr Thr Ser Gly Ala Met Leu Thr Gly Glu Leu Lys Lys Ala Leu Ile
 355 360 365

gag gtt ctg cag ccc ttg atc gca gag cac cag gcc cgg cgc aag gag 4573
 Glu Val Leu Gln Pro Leu Ile Ala Glu His Gln Ala Arg Arg Lys Glu
 370 375 380

gtc acg gat gag ata gtg aaa gag ttc atg act ccc cgg aag ctg tcc 4621
 Val Thr Asp Glu Ile Val Lys Glu Phe Met Thr Pro Arg Lys Leu Ser
 385 390 395

ttc gac ttt cag aag ctt gcg gcc gca ctc gag cac cac cac cac cac 4669
 Phe Asp Phe Gln Lys Leu Ala Ala Leu Glu His His His His His
 400 405 410

cac tgagatccgg ctgctaacaa agcccgaag gaagctgagt tggctgctgc 4722
 His
 415

caccgctgag caataactag cataaccct tggggcctct aaacgggtct tgagggggtt 4782

tttgctgaaa ggaggaacta tatccgat 4811

<210> 14

<211> 415

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: human
 supermini TrpRS in pET20B

<400> 14

Met Ser Asn His Gly Pro Asp Ala Thr Glu Ala Glu Glu Asp Phe Val
 1 5 10 15

Asp Pro Trp Thr Val Gln Thr Ser Ser Ala Lys Gly Ile Asp Tyr Asp
 20 25 30

Lys Leu Ile Val Arg Phe Gly Ser Ser Lys Ile Asp Lys Glu Leu Ile
 35 40 45

Asn Arg Ile Glu Arg Ala Thr Gly Gln Arg Pro His His Phe Leu Arg
 50 55 60

Arg Gly Ile Phe Phe Ser His Arg Asp Met Asn Gln Val Leu Asp Ala
 65 70 75 80
 Tyr Glu Asn Lys Lys Pro Phe Tyr Leu Tyr Thr Gly Arg Gly Pro Ser
 85 90 95
 Ser Glu Ala Met His Val Gly His Leu Ile Pro Phe Ile Phe Thr Lys
 100 105 110
 Trp Leu Gln Asp Val Phe Asn Val Pro Leu Val Ile Gln Met Thr Asp
 115 120 125
 Asp Glu Lys Tyr Leu Trp Lys Asp Leu Thr Leu Asp Gln Ala Tyr Gly
 130 135 140
 Asp Ala Val Glu Asn Ala Lys Asp Ile Ile Ala Cys Gly Phe Asp Ile
 145 150 155 160
 Asn Lys Thr Phe Ile Phe Ser Asp Leu Asp Tyr Met Gly Met Ser Ser
 165 170 175
 Gly Phe Tyr Lys Asn Val Val Lys Ile Gln Lys His Val Thr Phe Asn
 180 185 190
 Gln Val Lys Gly Ile Phe Gly Phe Thr Asp Ser Asp Cys Ile Gly Lys
 195 200 205
 Ile Ser Phe Pro Ala Ile Gln Ala Ala Pro Ser Phe Ser Asn Ser Phe
 210 215 220
 Pro Gln Ile Phe Arg Asp Arg Thr Asp Ile Gln Cys Leu Ile Pro Cys
 225 230 235 240
 Ala Ile Asp Gln Asp Pro Tyr Phe Arg Met Thr Arg Asp Val Ala Pro
 245 250 255
 Arg Ile Gly Tyr Pro Lys Pro Ala Leu Leu His Ser Thr Phe Phe Pro
 260 265 270
 Ala Leu Gln Gly Ala Gln Thr Lys Met Ser Ala Ser Asp Pro Asn Ser
 275 280 285
 Ser Ile Phe Leu Thr Asp Thr Ala Lys Gln Ile Lys Thr Lys Val Asn
 290 295 300
 Lys His Ala Phe Ser Gly Gly Arg Asp Thr Ile Glu Glu His Arg Gln
 305 310 315 320
 Phe Gly Gly Asn Cys Asp Val Asp Val Ser Phe Met Tyr Leu Thr Phe
 325 330 335
 Phe Leu Glu Asp Asp Asp Lys Leu Glu Gln Ile Arg Lys Asp Tyr Thr
 340 345 350
 Ser Gly Ala Met Leu Thr Gly Glu Leu Lys Lys Ala Leu Ile Glu Val
 355 360 365
 Leu Gln Pro Leu Ile Ala Glu His Gln Ala Arg Arg Lys Glu Val Thr
 370 375 380

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Asp Glu Ile Val Lys Glu Phe Met Thr Pro Arg Lys Leu Ser Phe Asp
 385 390 395 400

Phe Gln Lys Leu Ala Ala Ala Leu Glu His His His His His His
 405 410 415

<210> 15
 <211> 4742
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> CDS
 <222> (3428)..(4603)

<220>
 <223> Description of Artificial Sequence: human minor
 Trp-RS fragment in pET20B

<400> 15
 tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60
 cagcgtgacc gctacacttg ccagcgccct agcgcccgt cctttcgctt tcttcccttc 120
 ctttctcgcc acgttcgcgc gctttccccg tcaagctcta aatcgggggc tccctttagg 180
 gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc 240
 acgtagtggg ccacgcgcct gatagacggt ttttcgcctt ttgacgttgg agtccacgtt 300
 ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggctctattc 360
 ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 420
 acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480
 tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540
 tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat 600
 gagtattcaa catttccttg tcgcccttat tccctttttt gcggcatttt gccttctctg 660
 ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg 720
 agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt ttcgccccga 780
 agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg 840
 tattgacgcc gggcaagagc aactcggctg ccgcatacac tattctcaga atgacttggg 900
 tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg 960
 cagtgtgcc ataaccatga gtgataaac tgcgccaac ttacttctga caacgatcgg 1020
 aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga 1080
 tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc 1140
 tgcagcaatg gcaacaacgt tgcgcaaact attactggc gaactactta ctctagcttc 1200

ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc 1260
ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg 1320
cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac 1380
gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc 1440
actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt 1500
aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac 1560
caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa 1620
aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc 1680
accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt 1740
aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg 1800
ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa tcctgttacc 1860
agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt 1920
accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc ccagcttgga 1980
gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct 2040
tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggctcgaa caggagagcg 2100
cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca 2160
cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa 2220
cgccagcaac gcggcctttt tacggttctt ggccttttgc tggccttttg ctccatggtt 2280
ctttcctgcg ttatcccttg attctgtgga taaccgtatt accgcctttg agtgagctga 2340
taccgctcgc cgcagccgaa cgaccgagcg cagcaggtca gtgagcgagg aagcggaaga 2400
gcgcctgatg cggatatttc tccttacgca tctgtgcggt atttcacacc gcatatatgg 2460
tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagtatac actccgctat 2520
cgctacgtga ctgggtcatg gctgcgcccc gacaccgcc aacaccgct gacgcgcctt 2580
gacgggcttg tctgctcccg gcatccgctt acagacaagc tgtgaccgtc tccgggagct 2640
gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc gaggcagctg cggtaaagct 2700
catcagcgtg gtcgtgaagc gattcacaga tgtctgcctg ttcatccgcy tccagctcgt 2760
tgagtttctc cagaagcggt aatgtctggc ttctgataaa gcggggccatg ttaagggcgg 2820
ttttttcctg tttggctact gatgcctccg tgtaaggggg atttctgttc atgggggtaa 2880
tgataccgat gaaacgagag aggatgctca cgatacgggt tactgatgat gaacatgccc 2940
ggttactgga acgttgtgag ggtaaacaac tggcggtatg gatgcggcgg gaccagagaa 3000
aatcactca ggggtcaatgc cagcgcttcg ttaatacaga tgtaggtgtt ccacagggta 3060

gccagcagca	tcttcgcatg	cagatccgga	acataatggt	gcaggggcgct	gacttccgcg	3120										
tttccagact	ttacgaaaca	cggaaaccga	agaccattca	tgttggtgct	caggtcgcag	3180										
acgttttgca	gcagcagtcg	cttcacgttc	gctcgcgtat	cggtgattca	ttctgctaac	3240										
cagtaaggca	accccgccag	cctagccggg	tcctcaacga	caggagcacg	atcatgcgca	3300										
cccgtggcca	ggaccaaacg	ctgcccgaga	tctcgatccc	gcgaaattaa	tacgactcac	3360										
tatagggaga	ccacaacggt	ttccctctag	aaataatttt	gtttaacttt	aagaaggaga	3420										
tatacat	atg	agt	gca	aaa	ggc	ata	gac	tac	gat	aag	ctc	att	gtt	cgg	3469	
	Met	Ser	Ala	Lys	Gly	Ile	Asp	Tyr	Asp	Lys	Leu	Ile	Val	Arg		
	1				5					10						
ttt	gga	agt	agt	aaa	att	gac	aaa	gag	cta	ata	aac	cga	ata	gag	aga	3517
Phe	Gly	Ser	Ser	Lys	Ile	Asp	Lys	Glu	Leu	Ile	Asn	Arg	Ile	Glu	Arg	
15					20					25					30	
gcc	acc	ggc	caa	aga	cca	cac	cac	ttc	ctg	cgc	aga	ggc	atc	ttc	ttc	3565
Ala	Thr	Gly	Gln	Arg	Pro	His	His	Phe	Leu	Arg	Arg	Gly	Ile	Phe	Phe	
				35					40					45		
tca	cac	aga	gat	atg	aat	cag	gtt	ctt	gat	gcc	tat	gaa	aat	aag	aag	3613
Ser	His	Arg	Asp	Met	Asn	Gln	Val	Leu	Asp	Ala	Tyr	Glu	Asn	Lys	Lys	
				50				55					60			
cca	ttt	tat	ctg	tac	acg	ggc	cgg	ggc	ccc	tct	tct	gaa	gca	atg	cat	3661
Pro	Phe	Tyr	Leu	Tyr	Thr	Gly	Arg	Gly	Pro	Ser	Ser	Glu	Ala	Met	His	
		65					70					75				
gta	ggt	cac	ctc	att	cca	ttt	att	ttc	aca	aag	tgg	ctc	cag	gat	gta	3709
Val	Gly	His	Leu	Ile	Pro	Phe	Ile	Phe	Thr	Lys	Trp	Leu	Gln	Asp	Val	
	80					85					90					
ttt	aac	gtg	ccc	ttg	gtc	atc	cag	atg	acg	gat	gac	gag	aag	tat	ctg	3757
Phe	Asn	Val	Pro	Leu	Val	Ile	Gln	Met	Thr	Asp	Asp	Glu	Lys	Tyr	Leu	
95					100					105					110	
tgg	aag	gac	ctg	acc	ctg	gac	cag	gcc	tat	ggc	gat	gct	gtt	gag	aat	3805
Trp	Lys	Asp	Leu	Thr	Leu	Asp	Gln	Ala	Tyr	Gly	Asp	Ala	Val	Glu	Asn	
				115					120					125		
gcc	aag	gac	atc	atc	gcc	tgt	ggc	ttt	gac	atc	aac	aag	act	ttc	ata	3853
Ala	Lys	Asp	Ile	Ile	Ala	Cys	Gly	Phe	Asp	Ile	Asn	Lys	Thr	Phe	Ile	
			130					135					140			
ttc	tct	gac	ctg	gac	tac	atg	ggg	atg	agc	tca	ggg	ttc	tac	aaa	aat	3901
Phe	Ser	Asp	Leu	Asp	Tyr	Met	Gly	Met	Ser	Ser	Gly	Phe	Tyr	Lys	Asn	
		145					150					155				
gtg	gtg	aag	att	caa	aag	cat	gtt	acc	ttc	aac	caa	gtg	aaa	ggc	att	3949
Val	Val	Lys	Ile	Gln	Lys	His	Val	Thr	Phe	Asn	Gln	Val	Lys	Gly	Ile	
	160					165					170					
ttc	ggc	ttc	act	gac	agc	gac	tgc	att	ggg	aag	atc	agt	ttt	cct	gcc	3997
Phe	Gly	Phe	Thr	Asp	Ser	Asp	Cys	Ile	Gly	Lys	Ile	Ser	Phe	Pro	Ala	
175					180					185					190	

atc cag gct gct ccc tcc ttc agc aac tca ttc cca cag atc ttc cga	4045
Ile Gln Ala Ala Pro Ser Phe Ser Asn Ser Phe Pro Gln Ile Phe Arg	
195 200 205	
gac agg acg gat atc cag tgc ctt atc cca tgt gcc att gac cag gat	4093
Asp Arg Thr Asp Ile Gln Cys Leu Ile Pro Cys Ala Ile Asp Gln Asp	
210 215 220	
cct tac ttt aga atg aca agg gac gtc gcc ccc agg atc ggc tat cct	4141
Pro Tyr Phe Arg Met Thr Arg Asp Val Ala Pro Arg Ile Gly Tyr Pro	
225 230 235	
aaa cca gcc ctg ttg cac tcc acc ttc ttc cca gcc ctg cag ggc gcc	4189
Lys Pro Ala Leu Leu His Ser Thr Phe Phe Pro Ala Leu Gln Gly Ala	
240 245 250	
cag acc aaa atg agt gcc agc gac cca aac tcc tcc atc ttc ctc acc	4237
Gln Thr Lys Met Ser Ala Ser Asp Pro Asn Ser Ser Ile Phe Leu Thr	
255 260 265 270	
gac acg gcc aag cag atc aaa acc aag gtc aat aag cat gcg ttt tct	4285
Asp Thr Ala Lys Gln Ile Lys Thr Lys Val Asn Lys His Ala Phe Ser	
275 280 285	
gga ggg aga gac acc atc gag gag cac agg cag ttt ggg ggc aac tgt	4333
Gly Gly Arg Asp Thr Ile Glu Glu His Arg Gln Phe Gly Gly Asn Cys	
290 295 300	
gat gtg gac gtg tct ttc atg tac ctg acc ttc ttc ctc gag gac gac	4381
Asp Val Asp Val Ser Phe Met Tyr Leu Thr Phe Phe Leu Glu Asp Asp	
305 310 315	
gac aag ctc gag cag atc agg aag gat tac acc agc gga gcc atg ctc	4429
Asp Lys Leu Glu Gln Ile Arg Lys Asp Tyr Thr Ser Gly Ala Met Leu	
320 325 330	
acc ggt gag ctc aag aag gca ctc ata gag gtt ctg cag ccc ttg atc	4477
Thr Gly Glu Leu Lys Lys Ala Leu Ile Glu Val Leu Gln Pro Leu Ile	
335 340 345 350	
gca gag cac cag gcc cgg cgc aag gag gtc acg gat gag ata gtg aaa	4525
Ala Glu His Gln Ala Arg Arg Lys Glu Val Thr Asp Glu Ile Val Lys	
355 360 365	
gag ttc atg act ccc cgg aag ctg tcc ttc gac ttt cag aag ctt gcg	4573
Glu Phe Met Thr Pro Arg Lys Leu Ser Phe Asp Phe Gln Lys Leu Ala	
370 375 380	
gcc gca ctc gag cac cac cac cac cac cac tgagatccgg ctgctaacia	4623
Ala Ala Leu Glu His His His His His His	
385 390	
agcccgaaag gaagctgagt tggctgctgc caccgctgag caataactag cataaccct	4683
tggggcctct aaacgggtct tgaggggttt tttgctgaaa ggaggaacta tatccggat	4742

<210> 16
 <211> 392
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: human minor
 TrpRS fragment in pET20B

<400> 16

Met	Ser	Ala	Lys	Gly	Ile	Asp	Tyr	Asp	Lys	Leu	Ile	Val	Arg	Phe	Gly	1	5	10	15
Ser	Ser	Lys	Ile	Asp	Lys	Glu	Leu	Ile	Asn	Arg	Ile	Glu	Arg	Ala	Thr	20	25	30	
Gly	Gln	Arg	Pro	His	His	Phe	Leu	Arg	Arg	Gly	Ile	Phe	Phe	Ser	His	35	40	45	
Arg	Asp	Met	Asn	Gln	Val	Leu	Asp	Ala	Tyr	Glu	Asn	Lys	Lys	Pro	Phe	50	55	60	
Tyr	Leu	Tyr	Thr	Gly	Arg	Gly	Pro	Ser	Ser	Glu	Ala	Met	His	Val	Gly	65	70	75	80
His	Leu	Ile	Pro	Phe	Ile	Phe	Thr	Lys	Trp	Leu	Gln	Asp	Val	Phe	Asn	85	90	95	
Val	Pro	Leu	Val	Ile	Gln	Met	Thr	Asp	Asp	Glu	Lys	Tyr	Leu	Trp	Lys	100	105	110	
Asp	Leu	Thr	Leu	Asp	Gln	Ala	Tyr	Gly	Asp	Ala	Val	Glu	Asn	Ala	Lys	115	120	125	
Asp	Ile	Ile	Ala	Cys	Gly	Phe	Asp	Ile	Asn	Lys	Thr	Phe	Ile	Phe	Ser	130	135	140	
Asp	Leu	Asp	Tyr	Met	Gly	Met	Ser	Ser	Gly	Phe	Tyr	Lys	Asn	Val	Val	145	150	155	160
Lys	Ile	Gln	Lys	His	Val	Thr	Phe	Asn	Gln	Val	Lys	Gly	Ile	Phe	Gly	165	170	175	
Phe	Thr	Asp	Ser	Asp	Cys	Ile	Gly	Lys	Ile	Ser	Phe	Pro	Ala	Ile	Gln	180	185	190	
Ala	Ala	Pro	Ser	Phe	Ser	Asn	Ser	Phe	Pro	Gln	Ile	Phe	Arg	Asp	Arg	195	200	205	
Thr	Asp	Ile	Gln	Cys	Leu	Ile	Pro	Cys	Ala	Ile	Asp	Gln	Asp	Pro	Tyr	210	215	220	
Phe	Arg	Met	Thr	Arg	Asp	Val	Ala	Pro	Arg	Ile	Gly	Tyr	Pro	Lys	Pro	225	230	235	240
Ala	Leu	Leu	His	Ser	Thr	Phe	Phe	Pro	Ala	Leu	Gln	Gly	Ala	Gln	Thr	245	250	255	
Lys	Met	Ser	Ala	Ser	Asp	Pro	Asn	Ser	Ser	Ile	Phe	Leu	Thr	Asp	Thr	260	265	270	

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Ala Lys Gln Ile Lys Thr Lys Val Asn Lys His Ala Phe Ser Gly Gly
275 280 285

Arg Asp Thr Ile Glu Glu His Arg Gln Phe Gly Gly Asn Cys Asp Val
290 295 300

Asp Val Ser Phe Met Tyr Leu Thr Phe Phe Leu Glu Asp Asp Asp Lys
305 310 315 320

Leu Glu Gln Ile Arg Lys Asp Tyr Thr Ser Gly Ala Met Leu Thr Gly
325 330 335

Glu Leu Lys Lys Ala Leu Ile Glu Val Leu Gln Pro Leu Ile Ala Glu
340 345 350

His Gln Ala Arg Arg Lys Glu Val Thr Asp Glu Ile Val Lys Glu Phe
355 360 365

Met Thr Pro Arg Lys Leu Ser Phe Asp Phe Gln Lys Leu Ala Ala Ala
370 375 380

Leu Glu His His His His His His
385 390

<210> 17

<211> 6

<212> PRT

<213> Homo sapiens

<400> 17

Glu Leu Arg Val Ser Tyr
1 5

<210> 18

<211> 6

<212> PRT

<213> Escherichia coli

<400> 18

Glu Thr Val Gln Glu Trp
1 5

<210> 19

<211> 9

<212> PRT

<213> Homo sapiens

<400> 19

Ser Ala Lys Glu Leu Arg Cys Gln Cys
1 5

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

<210> 20
 <211> 11
 <212> PRT
 <213> Homo sapiens

<400> 20
 Ala Ser Val Ala Thr Glu Leu Arg Cys Gln Cys
 1 5 10

<210> 21
 <211> 7
 <212> PRT
 <213> Homo sapiens

<400> 21
 Ala Glu Leu Arg Cys Gln Cys
 1 5

<210> 22
 <211> 58
 <212> PRT
 <213> Homo sapiens

<400> 22
 Gly Asp Glu Lys Lys Ala Lys Glu Lys Ile Glu Lys Lys Gly Glu Lys
 1 5 10 15
 Lys Glu Lys Lys Gln Gln Ser Ile Ala Gly Ser Ala Asp Ser Lys Pro
 20 25 30
 Ile Asp Val Ser Arg Leu Asp Leu Arg Ile Gly Cys Ile Ile Thr Ala
 35 40 45
 Arg Lys His Pro Asp Ala Asp Ser Leu Tyr
 50 55

<210> 23
 <211> 58
 <212> PRT
 <213> Homo sapiens

<400> 23
 Pro Ala Leu Lys Lys Leu Ala Ser Ala Ala Tyr Pro Asp Pro Ser Lys
 1 5 10 15
 Gln Lys Pro Met Ala Lys Gly Pro Ala Lys Asn Ser Glu Pro Glu Glu
 20 25 30
 Val Ile Pro Ser Arg Leu Asp Ile Arg Val Gly Lys Ile Ile Thr Val
 35 40 45
 Glu Lys His Pro Asp Ala Asp Ser Leu Tyr
 50 55

<210> 24
 <211> 7
 <212> PRT
 <213> Homo sapiens

<400> 24
Arg Val Gly Lys Ile Ile Thr
1 5

<210> 25
<211> 7
<212> PRT
<213> Homo sapiens

<400> 25
Arg Ile Gly Cys Ile Ile Thr
1 5

<210> 26
<211> 7
<212> PRT
<213> Homo sapiens

<400> 26
Arg Ile Gly Arg Ile Ile Thr
1 5

<210> 27
<211> 7
<212> PRT
<213> Caenorhabditis elegans

<400> 27
Arg Val Gly Arg Ile Ile Lys
1 5

<210> 28
<211> 7
<212> PRT
<213> Saccharomyces cerevisiae

<400> 28
Arg Val Gly Phe Ile Gln Lys
1 5

<210> 29
<211> 7
<212> PRT
<213> Bos taurus

<400> 29
Arg Val Gly Lys Val Ile Ser
1 5

<210> 30
<211> 7
<212> PRT
<213> Mus musculus

<400> 30

```
<210> 31
<211> 7
<212> PRT
<213> Mesocricetus auratus
```

```
<210> 32
<211> 7
<212> PRT
<213> Ovis aries
```

```
<210> 33
<211> 7
<212> PRT
<213> Calcarea sp.
```

```
<210> 34
<211> 7
<212> PRT
<213> A. aeolicus
```

```
<210> 35
<211> 7
<212> PRT
<213> Escherichia coli
```

```
<210> 36
<211> 7
<212> PRT
<213> Escherichia coli
```

```
<400> 36
Arg Val Ala Leu Ile Glu Asn
  1             5
```

<400> 37
Arg Val Ala Lys Val Leu Lys
1 5

<400> 38
Arg Val Ala Glu Val Ile Glu
1 5

```
<400> 39
Arg Val Ala Glu Val Val Gln
  1               5
```

```
<400> 40
Arg Val Ala Glu Val Leu Ala
  1             5
```

<400> 41
Val Gly Glu Val Val Glu
1 5

```
<400> 42
Ile Gly His Val Leu Glu
  1                               5
```

<210> 43
 <211> 6
 <212> PRT
 <213> Synechococcus sp.

<400> 43
 Val Gly Arg Val Leu Glu
 1 5

<210> 44
 <211> 6
 <212> PRT
 <213> Thermus thermophilus

<400> 44
 Phe Ala Arg Val Leu Glu
 1 5

<210> 45
 <211> 85
 <212> PRT
 <213> Homo sapiens

<400> 45
 Met Ser Tyr Lys Ala Ala Ala Gly Glu Asp Tyr Lys Ala Asp Cys Pro
 1 5 10 15
 Pro Gly Asn Pro Ala Pro Thr Ser Asn His Gly Pro Asp Ala Thr Glu
 20 25 30
 Ala Glu Glu Asp Phe Val Asp Pro Trp Thr Val Gln Thr Ser Ser Ala
 35 40 45
 Lys Gly Ile Asp Tyr Asp Lys Leu Ile Val Arg Phe Gly Ser Ser Lys
 50 55 60
 Ile Asp Lys Glu Leu Ile Asn Arg Ile Glu Arg Ala Thr Gly Gln Arg
 65 70 75 80
 Pro His His Phe Leu
 85

<210> 46
 <211> 85
 <212> PRT
 <213> Bos taurus

<400> 46
 Thr Ser Tyr Lys Ala Ala Thr Gly Glu Asp Tyr Lys Val Asp Cys Pro
 1 5 10 15
 Pro Gly Asp Pro Ala Pro Glu Ser Gly Glu Gly Leu Asp Ala Thr Glu
 20 25 30
 Ala Asp Glu Asp Phe Val Asp Pro Trp Thr Val Gln Thr Ser Ser Ala
 35 40 45
 Lys Gly Ile Asp Tyr Asp Lys Leu Ile Val Arg Phe Gly Ser Ser Lys
 50 55 60

Pro His Arg Phe Leu
85

<400> 47
Met Ser Tyr Lys Ala Ala Met Gly Glu Glu Tyr Lys Ala Gly Cys Pro
1 5 10 15

Pro Gly Asn Pro Thr Ala Gly Arg Asn Cys Asp Ser Asp Ala Thr Lys
20 25 30

Ala Ser Glu Asp Phe Val Asp Pro Trp Thr Val Arg Thr Ser Ser Ala
35 40 45

Lys Gly Ile Asp Tyr Asp Lys Leu Ile Val Gln Pro Gly Ser Ser Lys
50 55 60

Ile Asp Lys Glu Leu Ile Asn Arg Ile Glu Arg Ala Thr Gly Gln Arg
65 70 75 80

Pro His Arg Phe Leu
85

```
<210> 48
<211> 85
<212> PRT
<213> Oryctolagus cuniculus
```

<400> 48
 Thr Ser Tyr Lys Glu Ala Met Gly Glu Asp Tyr Lys Ala Asp Cys Pro
 1 5 10 15

Pro Gly Asn Ser Thr Pro Asp Ser His Gly Pro Asp Glu Ala Val Asp
20 25 30

Asp Lys Glu Asp Phe Val Asp Pro Trp Thr Val Arg Thr Ser Ser Ala
35 40 45

Lys Gly Ile Asp Tyr Asp Lys Leu Ile Val Gln Phe Gly Ser Ser Lys
50 55 60

Ile Asp Lys Glu Leu Val Asn Arg Ile Glu Arg Ala Thr Gly Gln Arg
65 70 75 80

Pro His Arg Phe Leu
85

```

<400> 49
Ile Ser Tyr Gln Gly Arg Ile Pro Tyr Pro Arg Pro Gly Thr Cys Pro
  1             5             10             15
Gly Gly Ala Phe Thr Pro Asn Met Arg Thr Thr Lys Glu Phe Pro Asp
          20             25             30
Asp Val Val Thr Phe Ile Arg Asn His Pro Leu Met Tyr Asn Ser Ile
          35             40             45
Tyr Pro Ile His Lys Arg Pro Leu Ile Val Arg Ile Gly Thr Asp Tyr
          50             55             60
Lys Tyr Thr Lys Ile Ala Val Asp Arg Val Asn Ala Ala Asp Gly Arg
          65             70             75             80
Tyr His Val Leu Phe Leu
          85

```

```

<400> 50
Ile Ser Tyr Gln Gly Arg Ile Pro Tyr Pro Arg Pro Gly Thr Cys Pro
  1          5          10          15
Gly Gly Ala Phe Thr Pro Asn Met Arg Thr Thr Lys Asp Phe Pro Asp
      20          25          30
Asp Val Val Thr Phe Ile Arg Asn His Pro Leu Met Tyr Asn Ser Ile
      35          40          45
Ser Pro Ile His Arg Arg Pro Leu Ile Val Arg Ile Gly Thr Asp Tyr
      50          55          60
Lys Tyr Thr Lys Ile Ala Val Asp Arg Val Asn Ala Ala Asp Gly Arg
      65          70          75          80
Tyr His Val Leu Phe Leu
      85

```

<400> 51
Ala Ala Ala Gly Glu Asp Tyr Lys Ala Asp Cys Pro Pro Gly Asn Pro
1 5 10 15
Ala Pro Thr Ser Asn His Gly Pro Asp Ala Thr Glu Ala Glu Glu Asp
20 25 30

```
<210> 52
<211> 46
<212> PRT
<213> Bos taurus
```

Ala Pro Glu Ser Gly Glu Gly Leu Asp Ala Thr Glu Ala Asp Glu Asp
20 25 30

```
<210> 53
<211> 46
<212> PRT
<213> Mus musculus
```

Thr Ala Gly Arg Asn Cys Asp Ser Asp Ala Thr Lys Ala Ser Glu Asp
20 25 30

```
<210> 54
<211> 46
<212> PRT
<213> Oryctolagus cuniculus
```

Thr Pro Asp Ser His Gly Pro Asp Glu Ala Val Asp Asp Lys Glu Asp
20 25 30

```
<210> 55
<211> 41
<212> PRT
<213> Mus musculus
```

<400> 55
Ala Phe Ala Gly Glu Asp Phe Lys Val Asp Ile Pro Glu Thr His Gly
1 5 10 15

Gly Glu Gly Thr Glu Asp Glu Ile Asp Asp Glu Tyr Glu Gly Asp Trp
20 25 30

Ser Asn Ser Ser Ser Ser Thr Ser Gly
35 40

<210> 56

<211> 5

<212> PRT

<213> Homo sapiens

<400> 56

Met Gly Asp Ala Pro
1 5

<210> 57

<211> 5

<212> PRT

<213> Homo sapiens

<400> 57

Ser Asn His Gly Pro
1 5

<210> 58

<211> 5

<212> PRT

<213> Homo sapiens

<400> 58

Ser Ala Lys Gly Ile
1 5

1

1